

Overall functional gene diversity of microbial communities in three full-scale activated sludge bioreactors

Yu Xia · Xiaohui Wang · Xianghua Wen · Kun Ding ·
Jizhong Zhou · Yunfeng Yang · Yu Zhang

Received: 10 December 2013 / Revised: 25 April 2014 / Accepted: 26 April 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract Understanding microbial community composition is thought to be crucial for improving process functioning and stabilities of full-scale activated sludge reactors in wastewater treatment plants (WWTPs). However, functional gene compositions of microbial communities within them have not been clearly elucidated. To gain a complete picture of microbial community, in this study, GeoChip 4.2 was used to profile the overall functional genes of three full-scale activated sludge bioreactors, the 16S rRNA gene diversities of which had been unveiled by 454-pyrosequencing in our previous investigation. Triplicate activated sludge samples from each system were analyzed, with the detection of 38,507 to 40,654 functional genes. A high similarity of 77.3–81.2 % shared functional genes was noted among the nine samples, verified by the high 16S rRNA gene similarity with shared operational

taxonomic units (OTUs) constituting 66.4–70.0 % of the detected sequences in each system. Correlation analyses showed that the abundances of a wide array of functional genes were associated with system performances. For example, the abundances of carbon degradation genes were strongly correlated to chemical oxygen demand (COD) removal efficiencies ($r=0.8697$, $P<0.01$). Lastly, we found that sludge retention time (SRT), influent total nitrogen concentrations (TN_{inf}), and dissolved oxygen (DO) concentrations were key environmental factors shaping the overall functional genes. Together, the results revealed vast functional gene diversity and some links between the functional gene compositions and microbe-mediated processes.

Keywords Activated sludge processes · Functional gene · Treatment efficiencies · Environmental factors

Electronic supplementary material The online version of this article (doi:10.1007/s00253-014-5791-7) contains supplementary material, which is available to authorized users.

Y. Xia · X. Wang · X. Wen (✉) · K. Ding · J. Zhou · Y. Yang
Environmental Simulation and Pollution Control State Key Joint
Laboratory, School of Environment, Tsinghua University,
100084 Beijing, People's Republic of China
e-mail: xhwen@tsinghua.edu.cn

J. Zhou
Institute for Environmental Genomics and Department of Botany and
Microbiology, University of Oklahoma, Norman, OK, USA

Y. Zhang
State Key Laboratory of Environmental Aquatic Chemistry,
Research Center for Eco-Environmental Sciences, Chinese Academy
of Sciences, P.O. Box 2871, 100085 Beijing, China

Present Address:

X. Wang
Department of Environmental Science and Engineering, Beijing
University of Chemical Technology, 100029 Beijing, China

Introduction

Wastewater treatment plants (WWTPs), applying at very large scales ($\sim 10^5$ m³/day), are probably the largest applications of bioprocess engineering. It is both technically and economically important to achieve great stabilities and the highest efficiencies of WWTPs, which largely depends on the composition and activity of their microbial community (Wagner et al. 2002). Most WWTPs applied activated sludge processes, such as oxidation ditch (OD) and anaerobic-anoxic-oxic (A²O). A clear illustration of microbial diversity and functions of activated sludge systems would help elucidate the role of microorganisms in pollutant degradation and transformation within them.

Microorganisms are mostly uncultivable and of vast diversity, and the detection and characterization of them are quite

challenging (He et al. 2007). Early culture-dependent methods are biased by the selection of species which obviously do not represent the real dominance structure (Wagner et al. 1994). Subsequent molecular ways target genomic DNA and do not involve cultivation biases. However, some molecular tools including clone library and fingerprinting techniques can obtain incomplete results as they do not capture the whole complexity of microbial communities. Therefore, high-throughput approaches such as Illumina, 454-pyrosequencing, PhyloChips, and GeoChips have become flourished recently, as they provide more complete insights into overall microbial communities.

Despite the microbiological methods used, plenty of the studies on microbial communities of activated sludge bioreactors have been conducted to profile taxonomic information based on 16S rRNA genes, using technical approaches of clone library (Wells et al. 2011; Yang et al. 2011), denaturing gradient gel electrophoresis (DGGE) (Yang et al. 2011), terminal restriction fragment length polymorphism (T-RELFP) (Wells et al. 2011; Yang et al. 2011), fluorescence in situ hybridization (FISH) (Mielczarek et al. 2012, 2013), Illumina (Albertsen et al. 2012), 454-pyrosequencing (Hu et al. 2012; Wang et al. 2012a), and PhyloChips (Xia et al. 2010). However, 16S rRNA gene-based ways may not be efficient in distinguishing some closely related but ecologically distinct groups of microorganisms as a consequence of the low evolutionary rate of 16S rRNA genes (Palys et al. 1997). Alternatively, the characterizations of protein-coding genes (functional genes) may be ideal for monitoring the status of wastewater treatment systems, as functional genes provide a better opportunity for distinguishing very closely related ecological populations (Palys et al. 1997). Therefore, understanding functional gene structures of microbial communities may provide a more complete and accurate picture of activated sludge systems than only focusing on the taxonomic information relying on 16S rRNA genes.

Although some of the previous investigations targeted certain microorganisms in activated sludge systems based on functional genes, such as ammonia-oxidizing bacteria and archaea (AOB and AOA) (Whang et al. 2009; Wang et al. 2011, 2012b; Gao et al. 2014), denitrifying bacteria (Srinandan et al. 2011), and sulfur-oxidizing bacteria (SOB) (Luo et al. 2011), to date, few studies profiling overall functional genes have been reported. As wastewater contains an array of organic compounds including a wide range of lipids, proteins, and carbohydrates, and inorganic materials like metals (Gray 2010), we expect that the composition of functional genes of activated sludge is complicated. This prompted us to carry out a study to profile the overall functional genes; thus, a high-throughput tool based on functional gene markers

is needed to this end. The GeoChip, a comprehensive functional gene array, is one of such tools, as it provides incomparable insight into many individual functional genes in the same sample sets (Wang et al. 2009).

Having been shown to be highly specific and sensitive to the targets and of good quantitative capability (He et al. 2007), GeoChips are proved to be powerful for analyzing functional diversity and particularly useful for providing direct linkages of microbial genes/populations to system processes and functions (He et al. 2007). Recently, GeoChip 3.0 and the GeoChip 4 series have been applied to the investigations of industrial WWTPs (Zhang et al. 2013) and municipal wastewater treatment systems (Sun et al. 2014; Wang et al. 2014), respectively. The studies successfully revealed some linkages between functional potential and environmental parameters. The latest GeoChip in 2011, GeoChip 4.2, containing approximately 82,000 probes and covering 141,995 coding sequences in microbial functional processes such as nitrogen, carbon, phosphorus cycling, metal resistance, and organic contaminant degradation (Tu et al. 2014), could provide a good coverage of functional genes.

In this study, three full-scale activated sludge systems with relatively stable treatment efficiencies were investigated. The systems were located in the same city and did not have geological differences. The wastewater characteristics and operational parameters differed, and they might lead to some differences of microbial community structures. With the 16S rRNA gene profiles unveiled in our previous study, functional gene composition, richness, and evenness of the systems were analyzed by GeoChip 4.2 to achieve a full understanding of their microbial populations. Earlier studies using the methods of T-RFLP and clone library indicated that the detected richness (numbers of operational taxonomic units (OTUs)) and dominant phyla and OTUs did not change much (Wells et al. 2011) and revealed relatively stable AOB structures (Wang et al. 2011) in full-scale wastewater treatment systems with relatively stable treatment efficiencies during a year of monitoring. Other studies applying FISH suggested that the relative abundances of filamentous bacteria (Mielczarek et al. 2012), and abundant polyphosphate-accumulating organisms (PAOs) (Mielczarek et al. 2013) showed minor changes in full-scale WWTPs with good treatment efficiencies during 3-year investigations. From this perspective, we considered that a single grab of sample from functionally stable systems could be effective, as it could at least reveal microbial richness and abundant groups. Moreover, replicates of each system were analyzed to avoid possible errors of the GeoChip analysis. The scientific questions we address are (1) whether and how microbial compositions vary among the three systems, (2) what are the major factors shaping overall functional gene structures, and (3) whether there are direct links between compositions of specific functional genes and pollutant removal efficiencies.

Materials and methods

Location description and sampling

The full-scale activated sludge systems (LFA, LFB, and BH) were located in two WWTPs in the same city of Guangdong Province of China, having no differences in geology. LFA and LFB lied in the same WWTP, treating domestic wastewater and with the same inoculation sludge. BH lied in the other one treating domestic wastewater mixed with a small amount of food industrial wastewater (approximately 5 %). LFA applied the treatment process of adsorption-biodegradation (AB) with the biological zone divided into anaerobic, anoxic, and oxic parts. LFB applied the process of T-type oxidation ditch (OD). BH applied the modified University of Cape Town (MUCT) process, which is similar to the A²O process with returned sludge into different parts. In spite of fluctuations in municipal wastewater characteristics and pH and seasonal temperature variations of the bioreactors, the treatment systems kept good removal efficiencies of chemical oxygen demand (COD), total nitrogen (TN), and total phosphorus (TP) through biological processes, except that the phosphorus removal of BH relied on both biological processes and chemical precipitation (the addition of polymerization aluminum chloride, PAC). The average data (monthly) of their influent and effluent characteristics were collected.

On 13 August, 2011, mixed-liquor suspended sludge (MLSS) samples were taken from the aeration tanks of each system. At each site, three 50-ml samples were collected. Each sample was dispensed into a 50-ml sterile Eppendorf tube and centrifuged at 14,000*g* for 10 min. The pellets were stored at −80 °C for analysis. Meanwhile, the pollutant concentrations of influents and effluents were measured using standard methods (Chinese 2002) through analyzing the 24-h composite samples of this day. Temperatures, pH, and DO concentrations of the aeration tanks were measured by probes (Thermo Orion, MA, USA). Data of their MLSS concentrations and sludge retention time (SRT) were collected.

DNA extraction and purification

Genomic DNA was extracted from the pellets of activated sludge samples through joint use of freezing and sodium dodecyl sulfate (SDS) for cell lysis (Zhou et al. 1996). The extracted products were then purified employing the Wizard® SV Genomic DNA Purification Kit (Promega, Madison, WI).

DNA amplification, labeling, and hybridization

DNA (1,000 ng) was labeled, purified, and dried as previously described (Lu et al. 2012). All labeled DNA was resuspended in a 10-μl hybridization solution as previously described (Wu et al. 2006) and was hybridized with GeoChip 4.2 on a MAUI

hybridization station (BioMicro, Salt Lake City, UT, USA) at 42 °C with 40 % formamide for 16 h. Microarrays were scanned by a ScanArray 5000 Microarray Analysis System (PerkinElmer, Wellesley, MA, USA) at 100 % laser power.

Statistical analysis

Signal intensities of the spots were measured with ImaGene 6.0 (Biodiscovery Inc., El Segundo, CA, USA). Low-quality spots were removed before further data analysis as previously reported (He et al. 2010). The genes detected in only one out of three replicates from the same system were removed. Intensities of the genes were then transformed to the natural logarithmic form and normalized with mean signal intensity as described previously (Wu et al. 2008).

Hierarchical cluster analyses of overall functional genes and genes from several categories and detrended correspondence analysis (DCA) were done with R code (<http://www.r-project.org/>) to unveil functional gene similarities. Canonical correspondence analysis (CCA) by R code was performed to evaluate the effects of environmental parameters on overall functional genes. To select attributes in CCA modeling, variation inflation factors (VIF) were applied to examine whether the variance of canonical coefficients was inflated by the presence of correlations with other attributes. The attributes with VIF values larger than 20 were considered to be affected by other attributes and consequently removed (Yang et al. 2014). The Mantel test by R code was done to test the correlations of functional gene compositions and environmental variables. The magnitudes of the correlations were decided by the *r* and *P* values. If the correlations are statistically significant (*P* < 0.05), they can be considered good or strong when the *r* values are equal to 0.68 (absolute value) or higher, and *r* values (absolute value) between 0.36 to 0.67 generally suggest modest or moderate correlations (Taylor 1990).

Detailed information of the detected genes and their signal intensities is available at the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo/, accession number GSE53744).

Sample collection, DNA extraction, 454-pyrosequencing analysis, and data processing

The sampling locations of pyrosequencing analysis were the same as those of the GeoChip 4.2 analysis. Activated sludge samples were collected on 22 December, 2010, approximately 8 months earlier than the sampling for GeoChip 4.2 analysis. Detailed information of sample collection, DNA extraction, 454-pyrosequencing analysis, and data processing is shown in our previous study (Wang et al. 2012a). The samples from LFA, LFB, and BH were represented as A2, A1, and B in that study, respectively. 16S rRNA gene sequences are available at

NCBI Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra/>, accession numbers SRR952788 and SRR954283).

Results

Bioreactor performance and operational conditions

BH received wastewater of slightly higher COD and higher TN (TN inf). Monthly average data of wastewater characteristics of the systems during January to September of 2011 were shown in Table S1 (Supplementary Material). Their effluent COD values (COD eff) maintained below 30 mg/l; their effluent TP concentrations (TP eff) were below 1.3 mg/l; and the TN removal efficiencies of LFA, LFB, and BH were 46.5–70.4, 74.3–87.1, and 69.7–75.7 %, respectively. Geographic locations, influent and effluent characteristics, and temperatures and pH of the aeration tanks of the sampling day of the three systems were shown in Table 1. Some operational parameters (DO, SRT, and MLSS concentrations) are also shown (Table 1). The DO concentration of LFA was the highest, and the SRT of LFB was the longest.

Overall functional gene diversity

A total of 38,507 to 40,654 genes were detected among the nine samples, including a variety of genes involved in 15 microbial functional processes, such as carbon cycling, nitrogen cycling, phosphorus cycling, and metal resistance. Each sample showed high functional gene diversity and evenness as indicated by the α -diversity indices (Table 2). The detected numbers and total signal intensities of the genes in each category did not show significant differences (Fig. S1, Supplementary Material).

Similarity of functional gene compositions

Most of the detected genes were present within all samples, the total signal intensities of which accounting for 77.3 to 81.2 % in each sample. Besides, any two of the nine samples had 83.1–92.1 % overlapped genes (Table 2), indicating a great functional gene similarity of the investigated systems. This was verified by the high 16S rRNA gene similarity of the systems revealed by 454-pyrosequencing, with shared OTUs constituting 66.4–70.0 % of the detected sequences and the Bray-Curtis coefficients of any two systems being approximately 0.60. Replicate samples from the same location shared more functional genes. Correspondingly, hierarchical cluster analysis of overall functional genes showed that the replicates of each system grouped (Fig. S2, Supplementary Material), which was verified by the DCA result (Fig. S3, Supplementary Material).

Table 1 Treatment processes and capacities, wastewater characteristics, and operating parameters of the three systems

System	Process	Longitude	Latitude	Capacity (1,000 m ³ /day)	COD _{Cr} inf ^a (mg/l)	TN inf ^a (mg/l)	NH ₄ ⁺ -N inf ^a (mg/l)	TP inf ^a (mg/l)	COD _{Cr} eff ^b (mg/l)	TN eff ^b (mg/l)	NH ₄ ⁺ eff ^b (mg/l)	TP eff ^b (mg/l)	Temperature ^c (°C)	DO ^c (mg/l)	pH ^c	SRT (day)	MLSS (mg/l)
LFA	AB	114.15	22.54	100	150.0±2.6	27.8±0.8	11.7±0.1	2.7±0.1	20.0±0.5	14.9±0.1	1.7±0.1	0.3±0	25.0±0.1	3.26±0.08	6.41±0.01	18	4,800
LFB	OD	114.15	22.54	250	150.0±2.6	27.8±0.8	11.7±0.1	2.7±0.1	15.7±0.1	8.4±0.1	0.6±0	0.4±0	24.6±0.1	1.27±0.06	6.84±0.02	20	4,700
BH	MUCT	114.10	22.53	180	330.1±2.5	54.5±0.5	35.6±0.2	4.1±0	18.4±0.1	14.6±0	3.9±0	0.1±0.1	24.7±0.2	2.10±0.07	6.41±0.05	15	4,600

AB the process of adsorption-biodegradation, OD the process of oxidation ditch, MUCT the process of the modified University of Cape Town

^a Concentrations of COD, TN, NH₄⁺-N, and TP of the influents (24-h composite samples)

^b Concentrations of COD, TN, NH₄⁺-N, and TP of the effluents (24-h composite samples)

^c Temperatures, DO concentrations, and pH values of the aeration tanks

Table 2 Diversity indices and overlapped genes of the nine samples investigated

	LFA1	LFA2	LFA3	LFB1	LFB2	LFB3	BH1	BH2	BH3
LFA1	0	91.9 %	92.1 %	86.8 %	87.3 %	87.3 %	85.6 %	83.7 %	85.6 %
LFA2		0	89.4	86.1 %	85.0 %	86.5 %	84.6 %	83.1 %	84.3 %
LFA3			0	85.5 %	86.3 %	86.1 %	84.9 %	83.4 %	84.7 %
LFB1				0	90.2 %	90.9 %	85.3 %	84.5 %	85.5 %
LFB2					0	92.1 %	86.1 %	84.5 %	86.4 %
LFB3						0	85.2 %	84.4 %	86.0 %
BH1							0	87.2 %	91.4 %
BH2								0	89.8 %
BH3									0
Richness ^a	39,613	38,507	38,584	39,875	40,376	40,647	39,558	38,874	40,654
H^b	10.58	10.55	10.56	10.59	10.60	10.61	10.58	10.56	10.61
$11/D^c$	39266	38174	38251	39535	40017	40281	39227	38519	40313
J^d	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996

LFA1, LFA2, and LFA3 are the triplicates from system LFA; LFB1, LFB2, and LFB3 are the triplicates from system LFB; BH1, BH2, and BH3 are the triplicates from system BH

^a Detected gene number

^b The Shannon-Weiner index (higher number represents higher diversity)

^c The Reciprocal of Simpson's index (higher number represents higher diversity)

^d The Pielou's evenness index

Genes involved in important functional processes

Carbon-cycling genes

A total of 5,220 carbon-cycling genes were detected, involved in methanogenesis and methane oxidation, carbon compound degradation, carbon fixation, and acetogenesis. Among them, carbon degradation genes were of the largest numbers (75.8–76.4 %) and the highest abundances (75.7–76.4 %).

Degradation genes of 3,994 of several polysaccharides including starch, chitin, lignin, cellulose, and hemicellulose genes were detected. LFB and BH showed higher signal intensities of the degradation genes of each polysaccharide than LFA (Fig. 1). Moreover, BH was of the largest number of carbon degradation genes, with 172 genes only present within BH, 116 genes unique to LFB, and 79 genes unique to LFA. The abundances of carbon degradation gene and COD inf values showed a strong correlation, as revealed by the Mantel test results ($r=0.7483$, $P<0.05$).

There were 223 detected genes involved in methane metabolism, the total signal intensities of which accounted for 3.8–4.1 % of those of carbon-cycling genes; 121 genes associated with methane production were detected, and 110 genes were from Archaea; 102 genes (*mmoX* and *pmoA*) participating in methane oxidation were detected, 65 of which were from uncultured microorganisms. Of the 37 genes with known affiliations, most were from *Gammaproteobacteria* and *Alphaproteobacteria*. Among them, the genes from the

families of *Methylococcaceae* (seven genes) and *Beijerinckiaceae* (six genes) were relatively abundant.

Nitrogen-cycling genes

Nitrogen-containing substances in wastewater are inorganic or organic. For biological removal of amino nitrogen and heterocyclic nitrogen compounds, their conversion to ammonia is the first step. Then ammonia must be nitrified, and the nitrate denitrified to yield nitrogen. Depending on the kind of nitrogen compounds present in wastewater, nitrogen removal

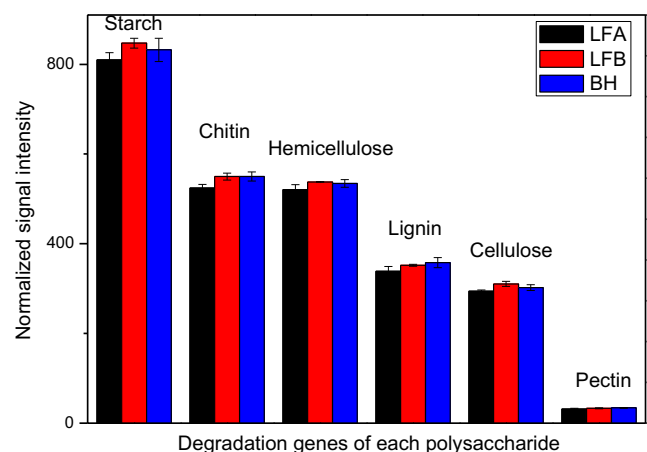


Fig. 1 Normalized signal intensities of the degradation genes of each polysaccharide

usually requires up to three processes in sequence: ammonification, nitrification, and denitrification. Of the detected nitrogen-cycling genes, the abundances of the genes participating in ammonification, nitrification, and denitrification were 12.1–12.5, 13.8–14.1, and 42.7–43.5 % in each sample, respectively. The total signal intensities of the genes associated with each process of nitrogen cycling and the numbers of the genes involved in important nitrogen-cycling bioprocesses in wastewater treatment were shown in Fig. S4 (Supplementary Material).

In ammonification, 466 genes (418 *ureC* and 48 *gdh* genes) were detected within all samples. Urea amidohydrolase (*ure*) encodes urease, serving to degrade amines (Wang and Marcotte 2008). The Mantel test results showed a significant correlation between TN inf values and ammonification gene compositions within all samples ($r=0.6546$, $P<0.01$).

In nitrification, 535 genes (512 *amoA* and 23 *hao* genes) were detected. Their total signal intensity of LFB was the highest, followed by that of BH (Fig. S4b, Supplementary Material). Besides, more genes were present within LFB and BH, with 9 genes unique to LFA, 20 genes only present within LFB, and 20 genes unique to BH. The unique nitrification genes of each system were shown in Fig. S5 (Supplementary Material). The Mantel test results revealed that nitrification gene abundances were moderately correlated to influent ammonia concentrations ($\text{NH}_4^+\text{-N}$ inf) ($r=0.6302$, $P<0.01$) and TN inf values ($r=0.622$, $P<0.01$) within all samples. The genes from the ammonia oxidizers which are often found in wastewater treatment systems were detected. Eleven genes were from *Nitrosomonas*, and three of them were from *Nitrosomonas eutropha* (114309057, 40809666, and 114332295), which are a model ammonia oxidizer (Wagner et al. 2002). Two genes were from *Nitrosococcus* (6117983 and 11640747). Eighteen genes were from *Nitrospira*, which have been reported to be present in WWTPs (Whang et al. 2009). The total signal intensity of the nitrification genes from *Nitrosomonas* of BH (8.61 ± 0.10) was slightly higher than that of LFA (8.52 ± 0.13) and LFB (8.06 ± 0.56), which was verified by the higher signal intensity of *gyrB* genes from *Nitrosomonas* within BH (5.17 ± 0.52) than that of LFA (3.62 ± 0.02) and LFB (4.28 ± 0.65). *GyrB* genes were used as a phylogenetic marker within GeoChips. Notably, 110 *amoA* genes from AOA were detected.

In denitrification, 1,597 genes were detected. Their total intensity within LFB was the highest, followed by that within BH (Fig. S4b, Supplementary Material). The Mantel test results revealed a significant correlation between the abundances of denitrification genes and TN inf concentrations within all samples ($r=0.7291$, $P<0.01$). As denitrifiers are phylogenetic diverse, the phylogenetic diversity of denitrifying bacteria was analyzed according to the GeoChip data. Most of them were from environmental clones rather than cultured isolates. Only 12.2 % genes were from cultured microorganisms. Among

them, *Proteobacteria* were predominant (10.7 %). The detected *Proteobacteria* consisted mainly of *Alphaproteobacteria* (48.0 %), *Gammaproteobacteria* (26.3 %), and *Betaproteobacteria* (22.2 %). In *Alphaproteobacteria*, 22 genes were from *Pseudomonas*, five genes were from *Paracoccus*, and four genes were from *Rhodobacter*. Strains of these genera have been isolated from activated sludge as efficient denitrifying bacteria (Magnusson et al. 1998).

Notably, eleven *hzo* genes encoding hydrazine oxidoreductase were detected, the total signal intensities of which constituted 0.2–0.3 % of those of nitrogen-cycling genes. This indicated the existence of anaerobic ammonium-oxidizing (anammox) bacteria. The detection of anammox genes in this investigation indicated that anammox is a process which may commonly exist in WWTPs.

Phosphorus-cycling genes

Phosphorus can lead to eutrophication as it would stimulate the excessive growth of phytoplankton in many surface waters. Therefore, the removal of phosphorus is substantially important. In total, 655 genes (*ppx*, *ppk*, and *phytase*) related to phosphorus cycling were detected, with 433 genes (66.1 %) present in all samples.

Polyphosphate kinase (*ppk*) is responsible for polyphosphate (polyP) synthesis; 241 *ppk* genes were present, 170 (70.5 %) of which were present in all samples (Fig. S6, Supplementary Material). For model bacteria, the *ppk* genes of which are highly conserved (Tzeng and Kornberg 1998): one gene from *Acinetobacter* sp. ADP1 (50084932), one gene from *Neisseria meningitidis* (254671991), one gene from *Deinococcus radiodurans* R1 (15807656), and one gene from *Vibrio cholerae* RC385 (254222220) were detected.

Exopolyphosphatase (*ppx*) is a highly processive enzyme catalyzing hydrolysis of terminal residues of long-chain polyP to phosphate (Pi) (Kornberg et al. 1999). A total of 382 *ppx* genes were detected; 247 (64.7 %) of them were present in all samples.

Relationships between community structures and environmental factors

To determine the effects of environmental variables, CCA was performed using six environmental variables (COD inf, $\text{NH}_4^+\text{-N}$ inf, TN inf, pH, DO, and SRT). The VIF values of the variables other than SRT, TN inf, and DO were over 20 and removed. SRT, TN inf, and DO were evaluated as key factors shaping the overall functional gene compositions (Fig. 2), which was confirmed by the Mantel test. The Mantel test results showed that overall functional gene compositions were strongly correlated to SRT ($r=0.7497$, $P<0.01$) and TN inf ($r=0.706$, $P<0.05$) and moderately correlated to DO concentrations ($r=0.5429$, $P<0.05$).

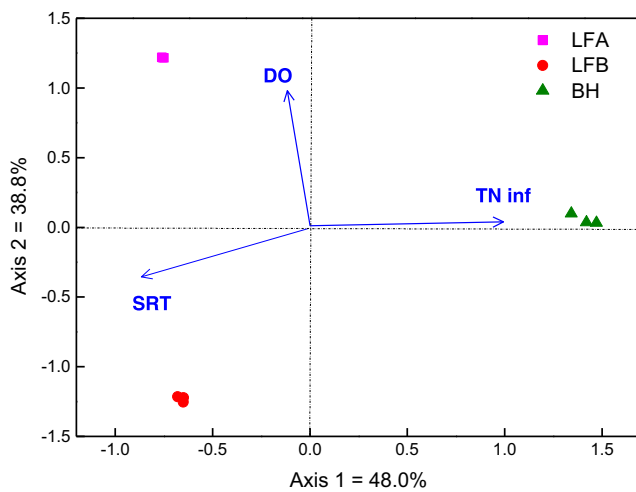


Fig. 2 Canonical correspondence analysis (CCA) of GeoChip 4.2 hybridization signal intensities and influent quality data as well as operational parameters. The amount of explained cumulative variation for axis 1 and axis 2 was 48.0 and 38.8 %, respectively

Relating removal efficiencies of COD and TN to corresponding functional gene structures

Pollutant removal rates are important indicators of WWTPs' efficiency. In this study, good biological removal rates of COD and TN were achieved within all systems (Table 3), which were determined by the equation of (influent substrate concentrations – effluent substrate concentrations)/influent substrate concentrations. The Mantel test results revealed some links between COD and TN removal rates and corresponding functional gene structures (Table 3). The compositions of carbon-cycling and carbon degradation genes were strongly associated with COD removal rates, and the abundances of the genes involved in nitrogen-cycling, ammonification, nitrification, and denitrification genes showed moderate correlations to TN removal rates. As not all of the systems achieved phosphorus removal relying merely on biological processes, phosphorus removal efficiencies were not directly

linked to phosphorus utilization gene structures when doing correlation analyses.

16S rRNA gene compositions revealed by 454-pyrosequencing

16S rRNA gene compositions of these systems had been unveiled by using 454-pyrosequencing (Wang et al. 2012a) investigating the samples collected 8 months earlier than those collected this time for GeoChip analysis, with 10,886 to 11,056 effective sequences detected within each system. The systems harbored highly similar microbial communities. At phylum level, *Proteobacteria* were the predominant group within all systems, accounting for 34.8–38.1 %, followed by *Bacteroidetes* (26.9–28.1 %) and *Chloroflexi* (14.2–17.2 %). As for some key groups, such as the genera of *Nitrosomonas* and *Nitrospira* which are important AOB known in WWTPs, they were of very low abundances within all the systems, accounting for 0.2 to 0.5 and 0 to 0.3 %, respectively.

Discussion

Compositions of microbial communities

This study showed a full picture of functional genes of three full-scale activated sludge systems. Vast and various functional genes were detected, which confirmed the powerfulness of GeoChip in analyzing activated sludge samples.

The compositions of the functional genes from certain categories detected in this study were consistent with those of an earlier study sequencing DNA and cDNA using Illumina Hi-seq2000. The dominance of denitrification genes in nitrogen-cycling genes was also reported in that study, which showed that denitrification genes sequences and their expression activities accounted for the largest in the processes of

Table 3 Relationship between functional gene signal intensities and substrate removal rates by the Mantel test

Pollutant	System	Removal efficiency (%)	Functional gene category	r_M^a	P
COD	LFA	86.7±0	Carbon cycling	0.8503	<0.01
	LFB	89.5±0.1	Carbon degradation	0.8697	<0.01
	BH	94.4±0.1			
TN	LFA	46.4±0.8	Nitrogen cycling	0.5261	<0.05
	LFB	69.8±0.4	Ammonification	0.4781	<0.05
			Nitrification	0.4663	<0.01
	BH	73.2±0	Denitrification	0.5211	<0.05

The signal intensity of different functional genes among nine samples was used as the first matrix; the corresponding substrate removal rates were used as the second matrix

^a r_M , Mantel's correlation coefficient

denitrification, ammonification, nitrogen fixation, and nitrification (Yu and Zhang 2012).

Similarity of microbial communities

High similarities of microbial communities of the investigated systems were detected, both in functional genes and 16S rRNA genes, consistent with some previous investigations. They have suggested that microbial populations showed a surprisingly consistent composition within wastewater treatment systems at different geographic locations (Xia et al. 2010; Wang et al. 2012a) and indicated higher microbial similarities (higher Jaccard's similarity coefficients) of the WWTPs in the same city (Wang et al. 2012a). Geographic location has been reported to be the most important factor shaping bacterial communities especially nonfecal groups in untreated sewage (Shanks et al. 2013). According to this, the influent bacterial community structures of the systems located in the same city may be quite similar, compared to those of geographically dispersed systems. Moreover, wastewater constituents and temperature of the systems in the same city are also similar. Therefore, it is reasonable that the investigated systems showed very high similarities of microbial communities.

Attributing functional gene differentiations to environmental factor variations

In spite of the high proportions of shared genes, functional gene structures of the investigated systems did show some differentiations, and we tried to link the differences to environmental factors by statistical analyses. The correlations between the concentrations of certain pollutants in influents and the abundances of corresponding functional genes suggested that wastewater characteristics did play a role in shaping the overall functional gene structures. Some details might be a reflection of this. COD_{inf} of BH was higher than that of the other two systems, and more carbon degradation genes were present within BH too. TN_{inf} and NH₄⁺-N_{inf} of BH were higher than those of LFA and LFB, and the nitrification genes and *gryB* genes from *Nitrosomonas*, which are suggested to be *r*-strategists (Andrews and Harris 1986), also showed slightly higher signal intensities within BH. Data of 454-pyrosequencing did not show such indications. The possible reason might be that *Nitrosomonas* might not be easily detected by the 454-pyrosequencing analysis due to their low abundances. Their comparisons among the investigated systems revealed by 454-pyrosequencing might be inaccurate. Besides, the total intensity of nitrification genes was slightly higher within BH than LFA, although LFA was of higher DO and longer SRT which might otherwise promote the nitrification process within it. Moreover, CCA results showed that SRT, TN_{inf}, and

DO were effective environmental factors shaping functional gene structures. Some discussions which might explain their importance are presented below.

SRT influences several processes in the activated sludge process such as nitrification. Longer SRT may favor the growth of slow-growing nitrifiers. The highest nitrification gene intensity within LFB might be due to its longest SRT, although LFA was of a higher DO concentration and BH was of higher TN_{inf} and NH₄⁺-N_{inf}. Besides, higher SRT allows longer contacting time between pollutants and bacteria, giving a better hydrolyzation and degradation of organic molecules (Massé et al. 2006). This might give an explanation to the result that LFB was of higher intensity of carbon degradation genes than LFA.

TN_{inf} is an important environmental variable for the growth and activity of the organisms involved in nitrogen cycling. Nitrogen loadings especially ammonia concentrations have been found to affect the compositions of AOB (Sui et al. 2014). With the MLSS concentrations being similar (Table 1), the nitrogen loadings of these systems mainly depend on the TN_{inf} values. As discussed above, the higher signal intensity and higher number of nitrification genes of BH than those of LFA and the slightly higher signal intensity of nitrification genes from *Nitrosomonas* of BH than that of the other two systems might result from the highest TN_{inf} of BH.

As with DO, many studies found that it did exert an effect on various categories of microbial groups in activated sludge, including AOB and nitrite-oxidizing bacteria (NOB) (Wang et al. 2012b; Liu and Wang 2013), denitrifying phosphorus-accumulating organisms (DPAOs), and denitrifying glycogen-accumulating organisms (DGAOs) (Yuan and Oleszkiewicz 2011), as well as heterotrophic bacteria (Wells et al. 2011). In the present study, DO was also evaluated as a key factor shaping the overall functional genes by CCA, and the highest signal intensity of denitrification genes within LFB might be due to the lowest DO value of its aeration tanks.

When performing CCA, treatment processes were not taken into account. There has been a great controversy on whether treatment processes play a role in shaping a microbial structure. An earlier study indicated that the AOB community of two different treatment processes—A²O and reversed A²O—treating identical water did not show significant differences (Wang et al. 2010). However, it was reported that the diversity of microbial community in samples from membrane bioreactors (MBRs) was the lowest compared with that of oxidation ditch and A²O/AO systems in another study (Hu et al. 2012). Whether the distinction of treatment processes impacts the microbial populations of a WWTP may depend on how much the treatment processes differ, and what kind of the microbes we focus on if the specific groups (e.g., AOB) or the overall microbial community. In this study, the applied three treatment processes all included anoxic and oxic parts and were similar to some extent.

Therefore, the treatment processes were thought to be similar and were not included when analyzing the effects of environmental variables on overall functional gene diversity in this study.

Relationship between functional gene structures and treatment efficiencies

The relationship between microbial community diversity and community functioning is an important issue in microbial ecology. In this study, some links between COD and TN treatment efficiencies and corresponding functional gene structures were detected by the Mantel test. The links were supported by some details as well. The highest number of carbon degradation genes was present within BH, and BH and LFB showed stronger signal intensities of carbon degradation genes than LFA. Correspondingly, the treatment efficiency in COD removal of BH was also the highest, followed by that of LFB. In addition, LFB and BH were of stronger signal intensities of nitrification genes and denitrification genes, and their TN removal efficiencies were higher too. The results are similar to a previous study to some extent, which indicated that denitrifying gene abundances could be used as a proxy for N₂O emission rates in soils (Morales et al. 2010). These links suggested that the highly diverse functional genes probably contributed to good treatment efficiencies of each system, and understanding functional gene compositions might help unveil treatment efficiencies of the investigated systems.

Acknowledgments This study was supported by the NSFC (51178239) and the special fund of the State Key Joint Laboratory of Environmental Simulation and Pollution Control (12L03ESPC).

References

- Albertsen M, Hansen LBS, Saunders AM, Nielsen PH, Nielsen KL (2012) A metagenome of a full-scale microbial community carrying out enhanced biological phosphorus removal. *ISME J* 6(6):1094–1106
- Andrews J, Harris R (1986) *r*- and *K*-Selection and microbial ecology. In: Marshall KC (ed) *Advances in microbial ecology*. Springer, New York, pp 99–147
- Chinese SEPA (2002) *Water and wastewater monitoring methods*, 4th edn. Chinese Environmental Science Publishing House, Beijing
- Gao J, Luo X, Wu G, Li T, Peng Y (2014) Abundance and diversity based on *amoA* genes of ammonia-oxidizing archaea and bacteria in ten wastewater treatment systems. *Appl Microbiol Biotechnol* 98(7):3339–3354
- Gray NF (2010) Chapter 13 - Nature of wastewater. In: Gray NF (ed) *Water technology*, 3rd edn. Butterworth-Heinemann, Oxford, pp 403–424
- He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC, Huang Z, Wu W, Gu B, Jardine P, Criddle C, Zhou J (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J* 1(1):67–77
- He Z, Xu M, Deng Y, Kang S, Kellogg L, Wu L, Van Nostrand JD, Hobbie SE, Reich PB, Zhou J (2010) Metagenomic analysis reveals a marked divergence in the structure of belowground microbial communities at elevated CO₂. *Ecol Lett* 13(5):564–575
- Hu M, Wang X, Wen X, Xia Y (2012) Microbial community structures in different wastewater treatment plants as revealed by 454-pyrosequencing analysis. *Bioresour Technol* 117:72–79
- Kornberg A, Rao NN, Ault-Riché D (1999) Inorganic polyphosphate: a molecule of many functions. *Annu Rev Biochem* 68:89–125
- Liu G, Wang J (2013) Long-term low DO enriches and shifts nitrifier community in activated sludge. *Environ Sci Technol* 47(10):5109–5117
- Lu Z, Deng Y, Van Nostrand JD, He Z, Voordeckers J, Zhou A, Lee Y-J, Mason OU, Dubinsky EA, Chavarria KL, Tom LM, Fortney JL, Lamendella R, Jansson JK, D'Haeseleer P, Hazen TC, Zhou J (2012) Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume. *ISME J* 6(2):451–460
- Luo JF, Lin WT, Guo Y (2011) Functional genes based analysis of sulfur-oxidizing bacteria community in sulfide removing bioreactor. *Appl Microbiol Biotechnol* 90(2):769–778
- Magnusson G, Edin H, Dalhammar G (1998) Characterisation of efficient denitrifying bacteria strains isolated from activated sludge by 16S-rDNA analysis. *Water Sci Technol* 38(8–9):63–68
- Massé A, Spérandio M, Cabassud C (2006) Comparison of sludge characteristics and performance of a submerged membrane bioreactor and an activated sludge process at high solids retention time. *Water Res* 40(12):2405–2415
- Mielczarek AT, Kragelund C, Eriksen PS, Nielsen PH (2012) Population dynamics of filamentous bacteria in Danish wastewater treatment plants with nutrient removal. *Water Res* 46(12):3781–3795
- Mielczarek AT, Nguyen HTT, Nielsen JL, Nielsen PH (2013) Population dynamics of bacteria involved in enhanced biological phosphorus removal in Danish wastewater treatment plants. *Water Res* 47(4):1529–1544
- Morales SE, Cosart T, Holben WE (2010) Bacterial gene abundances as indicators of greenhouse gas emission in soils. *ISME J* 4(6):799–808
- Palys T, Nakamura LK, Cohan FM (1997) Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int J Syst Bacteriol* 47(4):1145–1156
- Shanks OC, Newton RJ, Keltly CA, Huse SM, Sogin ML, McLellan SL (2013) Comparison of the microbial community structures of untreated wastewaters from different geographic locales. *Appl Environ Microbiol* 79(9):2906–2913
- Srinandan CS, Shah M, Patel B, Nerurkar AS (2011) Assessment of denitrifying bacterial composition in activated sludge. *Bioresour Technol* 102(20):9481–9489
- Sui Q, Liu C, Dong H, Zhu Z (2014) Effect of ammonium nitrogen concentration on the ammonia-oxidizing bacteria community in a membrane bioreactor for the treatment of anaerobically digested swine wastewater. *J Biosci Bioeng*. doi:10.1016/j.jbiosc.2014.02.017
- Sun Y, Shen YX, Liang P, Zhou J, Yang Y, Huang X (2014) Linkages between microbial functional potential and wastewater constituents in large-scale membrane bioreactors for municipal wastewater treatment. *Water Res* 56:162–171
- Taylor R (1990) Interpretation of the correlation coefficient: a basic review. *J Diagn Med Son* 6(1):35–39
- Tu Q, Yu H, He Z, Deng Y, Wu L, Van Nostrand JD, Zhou A, Voordeckers J, Lee YJ, Qin Y, Hemme CL, Shi Z, Xue K, Yuan T, Wang A, Zhou J (2014) GeoChip 4: a functional gene-array-based high-throughput environmental technology for microbial community analysis. *Mol Ecol Resour* 12(10):1755–1768
- Tzeng CM, Kornberg A (1998) Polyphosphate kinase is highly conserved in many bacterial pathogens. *Mol Microbiol* 29(1):381–382
- Wagner M, Erhart R, Manz W, Amann R, Lemmer H, Wedi D, Schleifer KH (1994) Development of an rRNA-targeted oligonucleotide probe specific for the genus *Acinetobacter* and its application for

- in situ monitoring in activated sludge. *Appl Environ Microbiol* 60(3):792–800
- Wagner M, Loy A, Nogueira R, Purkhold U, Lee N, Daims H (2002) Microbial community composition and function in wastewater treatment plants. *Antonie Van Leeuwenhoek* 81(1):665–680
- Wang F, Zhou H, Meng J, Peng X, Jiang L, Sun P, Zhang C, Van Nostrand JD, Deng Y, He Z, Wu L, Zhou J, Xiao X (2009) GeoChip-based analysis of metabolic diversity of microbial communities at the Juan de Fuca Ridge hydrothermal vent. *PNAS* 106(12):4840–4845
- Wang R, Marcotte EM (2008) The proteomic response of *Mycobacterium smegmatis* to anti-tuberculosis drugs suggests targeted pathways. *J Proteome Res* 7(3):855–865
- Wang X, Wen X, Criddle CS, Wells GF, Zhang J, Zhao Y (2010) Community analysis of ammonia-oxidizing bacteria in activated sludge of eight wastewater treatment systems. *J Environ Sci-China* 22(4):627–634
- Wang X, Wen X, Yan H, Ding K, Hu M (2011) Community dynamics of ammonia oxidizing bacteria in a full-scale wastewater treatment system with nitrification stability. *Front Environ Sci Eng* 5(1):92–98
- Wang X, Hu M, Xia Y, Wen X, Ding K (2012a) Pyrosequencing analysis of bacterial diversity in 14 wastewater treatment systems in China. *Appl Environ Microbiol* 78(19):7042–7047
- Wang X, Wen X, Xia Y, Hu M, Zhao F, Ding K (2012b) Ammonia oxidizing bacteria community dynamics in a pilot-scale wastewater treatment plant. *PLoS One* 7(4):e36272–e36272
- Wang X, Xia Y, Wen X, Yang Y, Zhou J (2014) Microbial Community functional structures in wastewater treatment plants as characterized by GeoChip. *PLoS One* 9(3):e93422
- Wells GF, Park H-D, Eggleston B, Francis CA, Criddle CS (2011) Fine-scale bacterial community dynamics and the taxa–time relationship within a full-scale activated sludge bioreactor. *Water Res* 45(17):5476–5488
- Whang L-M, Chien IC, Yuan S-L, Wu Y-J (2009) Nitrifying community structures and nitrification performance of full-scale municipal and swine wastewater treatment plants. *Chemosphere* 75(2):234–242
- Wu L, Liu X, Schadt CW, Zhou J (2006) Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl Environ Microbiol* 72(7):4931–4941
- Wu L, Kellogg L, Devol AH, Tiedje JM, Zhou J (2008) Microarray-based characterization of microbial community functional structure and heterogeneity in marine sediments from the Gulf of Mexico. *Appl Environ Microbiol* 74(14):4516–4529
- Xia S, Duan L, Song Y, Li J, Piceno Y, Andersen G, Alvarez-Cohen L, Moreno-Andrade I, Huang C, Hermanowicz S (2010) Bacterial community structure in geographically distributed biological wastewater treatment reactors. *Environ Sci Technol* 44(19):7391–7396
- Yang C, Zhang W, Liu RH, Li Q, Li BB, Wang SF, Song CJ, Qiao CL, Mulchandani A (2011) Phylogenetic diversity and metabolic potential of activated sludge microbial communities in full-scale wastewater treatment plants. *Environ Sci Technol* 45(17):7408–7415
- Yang Y, Gao Y, Wang S, Xu D, Yu H, Wu L, Lin Q, Hu Y, Li X, He Z, Deng Y, Zhou J (2014) The microbial gene diversity along an elevation gradient of the Tibetan grassland. *ISME J* 8(2):430–440
- Yu K, Zhang T (2012) Metagenomic and metatranscriptomic analysis of microbial community structure and gene expression of activated sludge. *PLoS One* 7(5):e38183
- Yuan QY, Oleszkiewicz JA (2011) Effect of dissolved oxygen on biological nutrient removal by denitrifying phosphorus-accumulating organisms in a continuous-flow system. *Water Environ Res* 83(11):2107–2114
- Zhang Y, Xie J, Liu M, Tian Z, He Z, van Nostrand JD, Ren L, Zhou J, Yang M (2013) Microbial community functional structure in response to antibiotics in pharmaceutical wastewater treatment systems. *Water Res* 47(16):6298–6308
- Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* 62(2):316–322