Available online at www.sciencedirect.com

ScienceDirect

www.journals.elsevier.com/journal-of-environmental-sciences

JOURNAL OF ENVIRONMENTAL SCIENCES
www.jesc.ac.cn

The first metagenome of activated sludge from full-scale anaerobic/anoxic/oxic (A2O) nitrogen and phosphorus removal reactor using Illumina sequencing

Mei Tian[1,2], *Fangqing Zhao*[3], *Xin Shen*[2], *Kahou Chu*[4], *Jinfeng Wang*[3], *Shuai Chen*[3], *Yan Guo*[5], *Hanhu Liu*[1,*]

1. *School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China*
2. *College of Marine Science, Huaihai Institute of Technology, Lianyungang 222005, China*
3. *Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China*
4. *School of Life Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong, China*
5. *Lianyungang Jinzhao Water Co., Ltd., Lianyungang 222005, China*

ARTICLE INFO

Article history:
Received 13 November 2014
Revised 1 December 2014
Accepted 9 December 2014
Available online 30 June 2015

Keywords:
Metagenome
Biodiversity
Anaerobic/anoxic/oxic (A2O)
Activated sludge
Nitrogen metabolism
ANAMMOX

ABSTRACT

The anaerobic/anoxic/oxic (A2O) process is globally one of the widely used biological sewage treatment processes. This is the first report of a metagenomic analysis using Illumina sequencing of full-scale A2O sludge from a municipal sewage treatment plant. With more than 530,000 clean reads from different taxa and metabolic categories, the metagenome results allow us to gain insight into the functioning of the biological community of the A2O sludge. There are 51 phyla and nearly 900 genera identified from the A2O activated sludge ecosystem. Proteobacteria, Bacteroidetes, Nitrospirae and Chloroflexi are predominant phyla in the activated sludge, suggesting that these organisms play key roles in the biodegradation processes in the A2O sewage treatment system. *Nitrospira*, *Thauera*, *Dechloromonas* and *Ignavibacterium*, which have abilities to metabolize nitrogen and aromatic compounds, are most prevalent genera. The percent of nitrogen and phosphorus metabolism in the A2O sludge is 2.72% and 1.48%, respectively. In the current A2O sludge, the proportion of *Candidatus* Accumulibacter is 1.37%, which is several times more than that reported in a recent study of A2O sludge. Among the four processes of nitrogen metabolism, denitrification related genes had the highest number of sequences (76.74%), followed by ammonification (15.77%), nitrogen fixation (3.88%) and nitrification (3.61%). In phylum Planctomycetes, four genera (*Planctomyces*, *Pirellula*, *Gemmata* and *Singulisphaera*) are included in the top 30 abundant genera, suggesting the key role of ANAMMOX in nitrogen metabolism in the A2O sludge.

Introduction

Water pollution is one of the major global environmental problems (Juma et al., 2014; Kroeze et al., 2013). Domestic sewage and industrial wastewater are increasingly affecting our precious freshwater resource. Among the many sewage treatment methods, the anaerobic/anoxic/oxic (A2O) process is a widely used biological sewage treatment process over the

* *Corresponding author*. E-mail: hanhucumt@sina.com (Hanhu Liu).

http://dx.doi.org/10.1016/j.jes.2014.12.027

past decades (Baek et al., 2012), which has an advantage of lower cost and energy requirement, and stable ability of nitrogen and phosphorus removal (Abu-Alhail and Lu, 2014). The A2O process is a sequential process that uses anaerobic, anoxic, and oxic reactors, and has been applied for removing nitrogen, phosphorous and organic carbons from sewage before discharge to receiving water (Kim et al., 2013). Even though many A2O sewage treatment plants run well, the phosphorus and nitrogen removal capacity can deteriorate sometimes, which often result from competition between beneficial and detrimental organisms in the treatment plants. So far, optimization of the A2O sewage treatment process has largely been based on empirical knowledge of process parameters and plant configuration. Hence, further understanding of the biological community and its functioning during treatment can lead to improved operation of A2O process.

It is estimated that there are at least 10^6–10^8 species of microorganisms in the world, but the vast majority (>99%) cannot be cultivated by conventional culturing methods (Liaw et al., 2010). Traditional metagenomic approach generally involves amplification of target gene fragments using PCR, following by building a library for sequencing. However, the problems of this approach at least include: 1) there are no "universal primers" for all taxa (including bacteria, archaea, fungi and virus) and therefore optimized PCR can only obtain part of the biodiversity information; and 2) PCR amplification efficiency may be biased toward a limited number of taxa.

Since 2005, high-throughput sequencing (HTS) technologies have been applied as novel promising methods to investigate biodiversity and metabolic function of various communities, including human guts (Karlsson et al., 2013; Qin et al., 2010; Ridaura et al., 2013; Schloissnig et al., 2013), oral cavity (Wang et al., 2013), ocean (DeLong, 2009; Metcalf et al., 2012; Walsh et al., 2009), and soils (Fierer et al., 2013; Guan et al., 2013; Mackelprang et al., 2011). However, metagenomic information of full-scale A2O nitrogen and phosphorus removal reactor in municipal sewage treatment plants is still very limited. The present study aims to investigate the broad spectrum of bacteria communities and metabolic functions in activated sludge from a typical municipal A2O reactor based on metagenomic approach and HTS technology. The metagenome data was used to investigate the biological communities in oxic sludge, and characterize the functional profiles of the biological community.

1. Materials and methods

1.1. Activated sludge sampling

Activated sludge was collected from the oxic reactor at the Dapu Sewage Treatment Plant in Lianyungang (LYG), Jiangsu province, China (34.647795° N, 119.168525° E; Table S1) on 10 Oct 2012. The plant using A2O treatment process has two biological pools, designed to treat a maximum of 250,000 m^3 of water per day. The size of each A2O biological pool is 88 m × 40 m × 7 m (L × B × H). The solid retention time (SRT) is 12 days, and the returned ratio is ranging from 73% to 84%. The detailed operational conditions of the wastewater treatment plants in that month can be found in the Table S2. The length of the sewage pipe network is more than 100 km, serving for 500,000 people. The plant has a total area of 191 acres with 13 pump stations, and is the largest urban sewage treatment plant in Lianyungang city and one of the key projects of the Huaihe River basin pollution prevention. The plant had stable A2O performance for several years. The fresh sludge sample collected was kept on ice and immediately transported to the laboratory. After centrifugation at 12,000 *g* for 20 min, the supernatant was discarded, and the sludge sample was stored at −76°C ultra-low temperature refrigerator.

1.2. Metagenomic DNA extraction and Illumina sequencing

DNA extraction was conducted within 24 h after sampling. DNA was extracted from sludge with QIAamp DNA Kit (Qiagen) following the instructions of the manufacturer. The quantity and quality of isolated DNA were evaluated using a Nano Drop spectrophotometer (Thermo) and 0.8% agarose gel electrophoresis (Bio-Rad), respectively.

Approximately 5 μg purified DNA sample was used for shotgun paired-end library construction. A library of ~180 bp DNA fragments was prepared according to a standard protocol of Illumina, and then sequenced by Illumina HiSeq 2000 platform at the Beijing Institutes of Life Science, Chinese Academy of Sciences (BIOLS, CAS). Base-calling was performed by the Illumina Genome Analyzer Pipeline software, and more than one million reads of raw sequences were generated. The HTS platforms often produce artificial replicates, and failure to remove these replicated sequences could lead to incorrect conclusion (Yang et al., 2014). Technical duplicate reads and the raw reads containing three or more undetermined bases or contaminated with adaptors were removed from the dataset by custom scripts (Guan et al., 2013). The remaining clean paired-end reads were merged into long reads (160–190 bp) for further analyses.

1.3. Taxonomic classification and functional assignment

These merged long read sequence sets were blasted for sequence matching using the NR (non-redundant) database, and the BLASTX alignments were further processed by the MEGAN 5 (Huson et al., 2011) to statistically analyze the abundance of each taxon. MEGAN (MEta Genome ANalyzer) software uses a homology-matching algorithm to generate a phylogenetic tree using the GenBank taxonomic database. For comparing the taxonomic composition and gene functions between different active sludge samples (Martin et al., 2006), two Enhanced Biological Phosphorus Removal (EBPR) sludge metagenomic data from USA and Australia (AU) (Table S1) were obtained from MG-RAST platform (http://metagenomics.anl.gov/) (Meyer et al., 2008). After normalizing the sequence counts of each taxon by the total number of reads, taxonomic analyses were performed on the biological composition and abundance at the domain, phylum and genus levels. The sequences from BLAST results were then assigned to GenBank taxonomies with MEGAN 5 using the Lowest Common Ancestor (LCA) algorithm, which assigns a sequence to the lowest common ancestor if it cannot be assigned uniquely to a given species.

For functional assignments, the metagenomic data were annotated against SEED subsystems in MEGAN 5 (Mitra et al.,

2011). The annotated sequences were sorted into 27 subsystems (Level 1) of SEED database to provide an overall profile of biological functions. To study further on nitrogen and phosphorus metabolism, which is closely related to nitrogen and phosphorus removal in the A2O sludge, Level 2 SEED subsystems of MEGAN was applied to annotate nitrogen and phosphorus metabolism related genes. In addition, the nitrogen removal related gene identified based on MEGAN results were extracted and applied to BLASTX against GenBank NR database, and then the BLAST results were visualized using the KEGG (Kyoto Encyclopedia of Genes and Genomes) mapper with the default parameters (Mitra et al., 2011).

2. Results and discussion

2.1. Biological community in the A2O active sludge

There are 51 phyla and nearly 900 genera identified in the A2O active sludge ecosystem, which is much higher than previous reports based on PCR-DGGE and FISH (Mao et al., 2008; Sun et al., 2013; Wu et al., 2011). Within the four domains, the majority of these sequences could be assigned to Bacteria (530,204 reads, 99.25%), and the remaining sequences were assigned to Archaea (2061 reads, 0.39%), Eukarya (1404 reads, 0.26%) and Virus (561 reads, 0.11%) (Fig. 1). The software generated two phylogenetic trees based on the GenBank taxonomic database (Fig. S1 & S2), in which the size of each circular node is proportional to the number of assignments at the particular taxonomic level. All phyla found in USA and AU sludge samples were detected in LYG sludge metagenome (Martin et al., 2006). However, nine groups detected in current metagenome were not found in previous two sludge samples, including Ignavibacteriae (10,883 reads), Candidatus Saccharibacteria (879 reads), Viruses (561 reads), Cloacimonetes (96 reads), Thermodesulfobacteria (33 reads), Caldiserica (21 reads), Chytridiomycota (12 reads), Latescibacteria (7 reads) and Aminicenantes (5 reads). Among them, Ignavibacteriae even ranked as the sixth largest phylum (accounting for 2.05% of Bacteria) in the 51 phyla of the A2O sludge metagenome. In addition, a considerable amount of sequences (218,551 reads) could not be taxonomically assigned, demonstrating the astonishing high biological diversity present in the A2O sludge. These unsigned sequences most likely represent uncharacterized organisms or novel gene fragments that are not related to or are very distantly related to any known sequences deposited in the current NR databases.

2.2. Bacteria community composition and abundance

The Bacteria domain was represented by 34 different phyla, with Proteobacteria, Bacteroidetes, Nitrospirae, Chloroflexi, and Planctomycetes ranking as the top five most abundant ones (Fig. 1). Proteobacteria was found to be the most abundant phylum (30,8575 reads, 58.20%), followed by Bacteroidetes (68,614 reads, 12.94%), Nitrospirae (63,859 reads, 12.04%) and Chloroflexi (22,509 reads, 4.25%). This finding supports that Proteobacteria, Bacteroidetes, Nitrospirae and Chloroflexi are predominant phyla in the active sludge, suggesting that these organisms play key roles in the biodegradation processes in A2O sewage treatment system. Other abundant phyla of Bacteria included Planctomycetes (15,918 reads, 3.00%), Ignavibacteria (10,883 reads, 2.05%), Actinobacteria (9286 reads, 1.75%), Firmicutes (8172 reads, 1.54%), Verrucomicrobia (6737 reads, 1.27%) and Acidobacteria (6331 reads, 1.19%) (Fig. 1).

There are 742 bacterial genera identified in the A2O sludge, which reveal a much higher bacterial diversity in A2O sewage treatment sludge as compared to previous reports on this kind of sludge (Bae et al., 2010; Kim et al., 2013). Among them, *Nitrospira* (63,658 reads, 23.92%), *Thauera* (16,882 reads, 6.34%), *Dechloromonas* (15,591 reads, 5.86%) and *Ignavibacterium* (10,883 reads, 4.09%) are the most prevalent. *Nitrospira*, the genus with the highest number of reads in the A2O sludge metagenome, accounts for almost a quarter of the Bacteria domain. In fact, *Nitrospira* bacteria are the key nitrite oxidizers in sewage treatment plants (Chiellini et al., 2013; Gilbert et al., 2014; Jogler et al., 2011; Kostan et al., 2010; Maixner et al., 2008; Ushiki et al., 2013). The second and third abundant genera, *Thauera* and *Dechloromonas*, have the ability to metabolize aromatic compounds (Coates et al., 2001; Liu et al., 2013; Mao et al., 2013; Salinero et al., 2009). The fourth ranked genus *Ignavibacterium* is a group of chemoheterotrophs with a versatile metabolism (Liu et al., 2012). Other abundant genera of Bacteria include *Caldilinea* (7616 reads, 2.86%), *Sorangium* (7226 reads, 2.72%), *Nitrosomonas* (6585 reads, 2.48%), *Anaerolinea* (5720 reads, 2.15%), *Niastella* (5493 reads, 2.06%), *Haliscomenobacter* (3944 reads, 1.48%), *Candidatus* Accumulibacter (3656 reads, 1.37%) and *Planctomyces* (3467 reads, 1.30%) (Fig. 2).

Nitrite-oxidizing bacteria *Nitrospira*, ammonia-oxidizing bacteria *Nitrosomonas* and phosphate-accumulating bacteria *Candidatus* Accumulibacter that have been found in the A2O process (Kim et al., 2013), were identified in the current metagenome (*Nitrospira*: 63,658 reads, 23.92%; *Nitrosomonas*: 6585 reads, 2.48%; *Candidatus* Accumulibacter: 3656 reads, 1.37%) (Fig. 2). The sequences of *Paracoccus* and *Alcaligenes* (well-known aerobic denitrifiers genera involved denitrification function) are 105 and 39, respectively. In addition, the genus *Thauera* is the second most abundant genus in the A2O metagenome (16,882 reads), accounting for 6.34% in the Bacteria domain. In a previous study, however, only 26 reads for the genus *Thauera* were detected out of a total of 7447 reads, and the genera *Paracoccus* and *Alcaligenes* were not detected at all (Kim et al., 2013). Many genera are not detected in the previous A2O sludge. The possible reasons are very low sequencing throughput (only 7447 reads in a total) and the distortion caused by PCR amplification. For example, the genus *Nitrosospira*, which is a well-known ammonia-oxidizing bacterium (Winkler et al., 2013; Zhang et al., 2014) was not observed in the previous study on A2O sludge (Kim et al., 2013). Yet in the present study the genus was detected with 353 reads and ranked as one of the top 100 most abundant genera in the A2O sludge metagenome (Fig. 2).

2.3. Archaea, Eukarya and Viruses community

In the domain Archaea, there are four phyla and 50 genera detected in the A2O sludge metagenome. Euryarchaeota was revealed as the most abundant phylum (1916 reads, 92.97%), followed by Crenarchaeota (83 reads, 4.03%), Thaumarchaeota (56 reads, 2.72%) and Korarchaeota (6 reads, 0.29%). At the

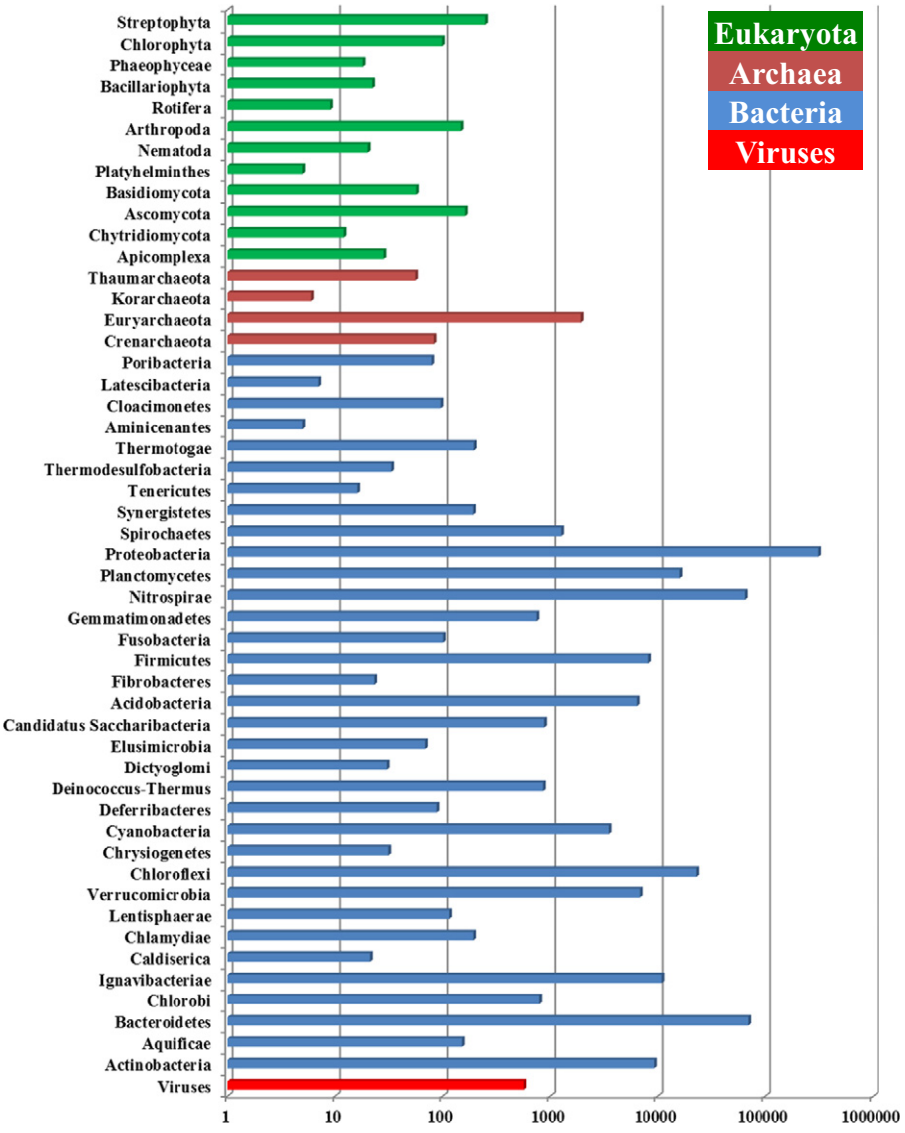

Fig. 1 – Biological composition and abundance (phylum level) in the A2O sludge metagenome. A2O: anaerobic/anoxic/oxic.

genus level, *Methanosaeta* is the largest genus (474 reads), accounting for 29.48% in the domain Archaea and is the only non-bacteria genus ranked as one of the top 100 most abundant genera (ranking 70th) (Fig. 2). *Methanosaeta* spp. are some of the most active methanogens and producing a considerable amount of methane on the planet. The second largest genus is *Methanosarcina* (148 reads, 9.20%), followed by *Methanospirillum* (105 reads, 6.53%). In addition, there are 77 eukaryotes genera detected in the A2O sludge ecosystem, which belong to 12 phyla. Among them, Streptophyta is the largest phylum (264 reads, 18.80%), followed by Ciliophora (242 reads, 17.24%) and Ascomycota (161 reads, 11.47%).

Like the virome in raw sewage (Cantalupo et al., 2011), viruses in the A2O sludge metagenome are dominated by bacteriophages. The majority of viruses in the A2O sludge belong to the order Caudovirales. Meanwhile, Siphoviridae, a family of double-stranded DNA viruses infecting only bacteria (Brussow and Desiere, 2001), is the largest family (116 reads) identified in the sludge. The characteristic structural features of this family are a nonenveloped head and noncontractile tail. Myoviridae (56 reads) and Podoviridae (33 reads) are the second and third largest families, respectively. Due to the limitation of the PCR-based method used in the previous study on A2O sludge (Kim et al., 2013), the diversity of viruses could not be revealed. The occurrence of the virus may have potential impact on the performance of the sewage treatment and receiving water body, and warrants further studies.

2.4. Abundance of functional genes in the A2O sludge

The annotation of functional genes was conducted in MEGAN 5 using SEED subsystems (Fig. 3). The phylogenetic richness observed reflected the wide metabolic diversity present in the A2O sludge metagenome. Among the Level 1 subsystems, the subsystem of carbohydrates was the most abundant (48,042 reads, 13.37%). The second largest subsystem is virulence, disease and defense (36,636 reads, 10.20%), followed by amino acids and derivatives (34,414 reads, 9.58%). In addition, other

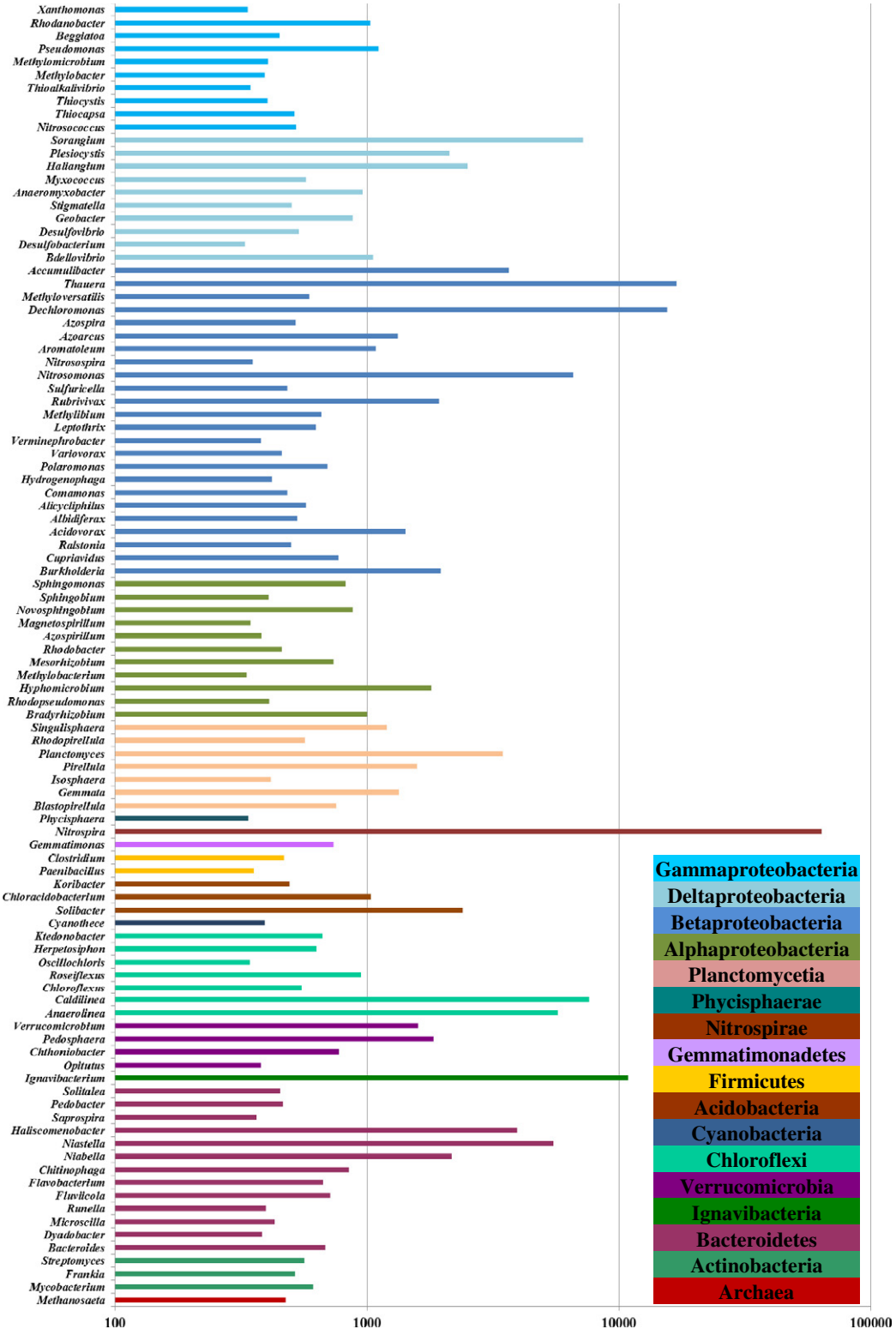

Fig. 2 – Top 100 most abundant genera in the A2O sludge metagenome. A2O: anaerobic/anoxic/oxic.

abundant subsystems are protein metabolism (29,316 reads, 8.16%), DNA metabolism (23,837 reads, 6.64%) and regulation and cell signaling (21,956 reads, 6.11%). Genes assigned to the metabolism of carbohydrates, amino acids and proteins result from the growth and production of living organisms occurred in high abundance in the treatment plant.

The presence of sequences assigned to functions like nitrogen, phosphorus and aromatic compounds metabolisms are essential

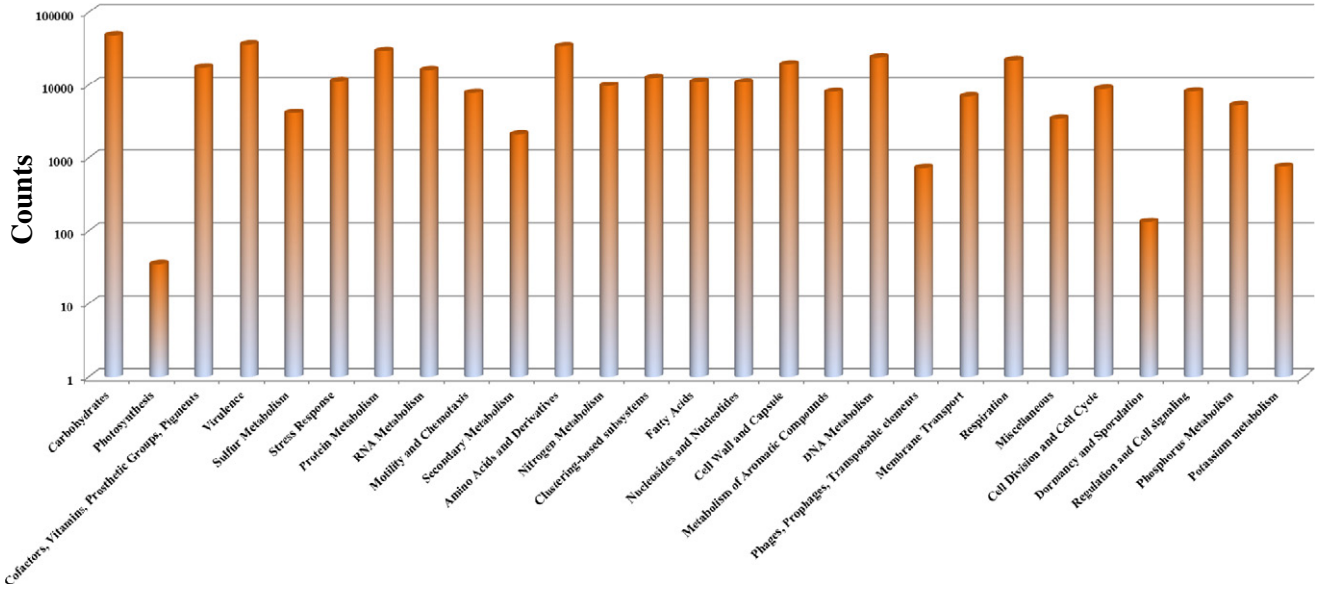

Fig. 3 – Abundance of SEED subsystems (Level 1) in the A2O sludge metagenome. A2O: anaerobic/anoxic/oxic.

for the performance of sewage treatment plant, since they are indicative that the microorganisms from the sludge are degrading and/or assimilating such compounds (Silva et al., 2012). The percent of nitrogen metabolism in the A2O sludge is 2.72% (9756 reads), which is higher than those in two EBPR treatment plants (1.94% and 1.64% for USA and AU sludge, respectively). Meanwhile, similar situation is also found in the metabolism of aromatic compounds. In the current A2O sludge, the proportion of this metabolism is 2.24% (8031 reads), but those in two EBPR samples are only 1.44% and 1.56% (for USA and AU sludge samples, respectively). In the USA and AU EBPR sludge, 1115 and 388 reads involved phosphate removal were found, and the proportions were 0.89% (Martin et al., 2006). In contrast, a total of 5331 gene sequences involved phosphorus

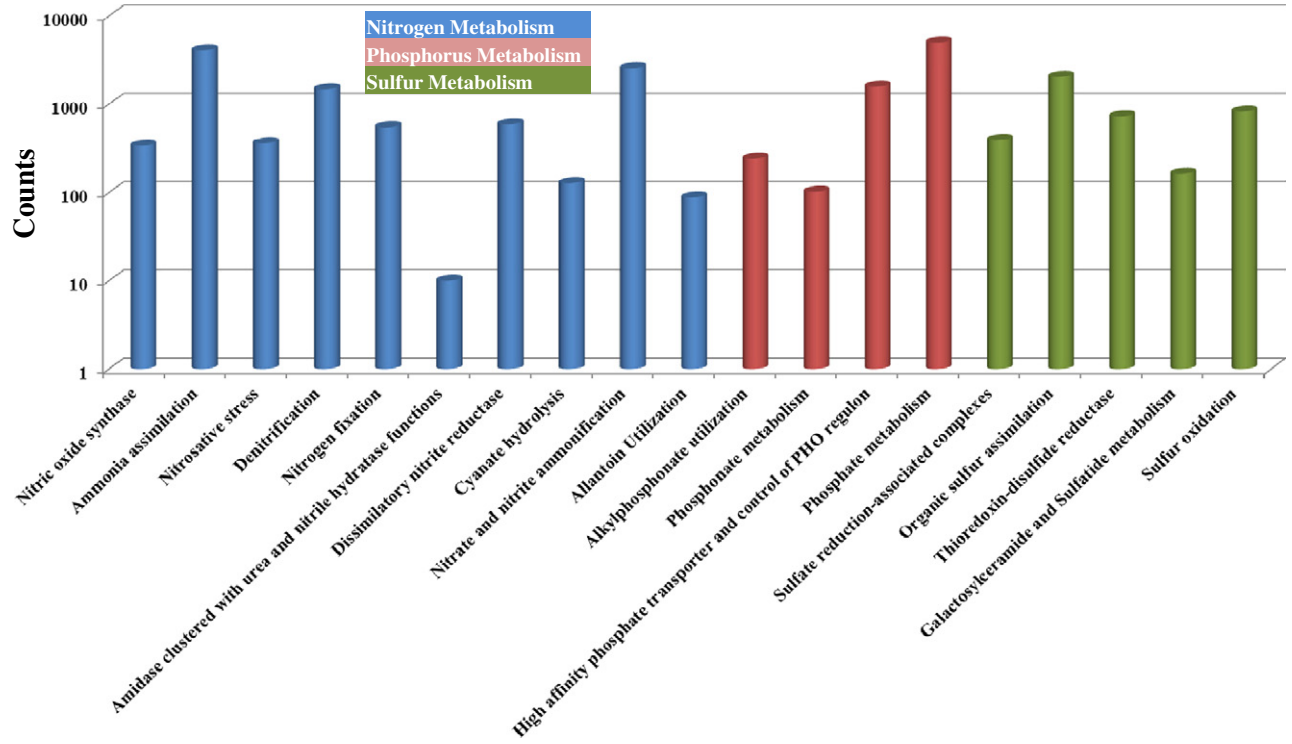

Fig. 4 – Nitrogen, phosphorus, and sulfur metabolism classification analysis based on Level 2 SEED subsystems.

metabolism were identified in the A2O metagenome (accounting for 1.48%).

Subsystems of nitrogen and phosphorus metabolism were subjected to further classification analyses (Level 2 SEED subsystems), as illustrated in Fig. 4. There are 9756 sequences assigned to the nitrogen metabolism in the A2O sludge, the majority of which represented genes coding enzymes linked to processes such as nitrification, denitrification, ammonification and nitrogen fixation, which are considered important processes for the removal of complex nitrogen compounds in sewage treatment. Ammonia assimilation and nitrite/nitrate ammonification-related genes had the highest sequence numbers (4024 and 2551, respectively), followed by denitrification (1464 sequences), dissimilatory nitrite reductase (589 sequences) and nitrogen fixation (547 sequences) (Fig. 4). Level 2 SEED subsystems of phosphorus and sulfur metabolism are also shown.

Microorganisms determine the function of a biological sewage treatment process for phosphate removal. The genus *Candidatus* Accumulibacter was reported to be directly responsible for phosphate accumulation (Martin et al., 2006); it is an unclassified group of Betaproteobacteria and a common bacterial community member of sewage treatment plants performing EBPR. Kim et al. (2013) found that this genus comprised 0.1%–0.5% of all bacteria in their A2O process (Kim et al., 2013). However, in the current A2O sludge, the proportion of *Candidatus* Accumulibacter is 1.37% (3656 reads) (Fig. 2), several times higher than in the above study. The underestimate of *Candidatus* Accumulibacter by Kim et al. (2013) may result from the deficiency of the PCR-based method. A high abundance of phosphate-accumulating bacteria *Candidatus* Accumulibacter can help improve the ability to store and use polyphosphate in the A2O sludge.

Besides using the SEED subsystems, we also conducted functional analysis of the A2O activated sludge metagenomic sequences using the classification and abundance of KEGG analysis (Fig. 5). All functional gene sequences can be divided into six categories and 36 subcategories. Among the six categories, the metabolism relevant functional genes are most abundant (114,582 sequences, 61.19%), followed by environmental information processes (28,734 sequences, 15.35%) and genetic information processing (28,347 sequences, 15.14%).

2.5. *Functional genes and pathways related nitrogen metabolism*

The phylogenetic richness observed reflects the wide metabolic diversity present in the metagenomic data from the A2O sludge (Fig. 3). Sequences associated with the four processes, *i.e.* nitrification, denitrification, ammonification, and nitrogen fixation, were extracted according to the BLAST results matched against NCBI-nr database and then mapped with KEGG analyzer. The sequence number of denitrification related genes is the highest (2633 reads, 76.74%), followed by ammonification (541 reads, 15.77%), nitrogen fixation (133 reads, 3.88%) and nitrification (124 reads, 3.61%). The dominance of denitrification coding gene sequences among the four processes is consistent with result from a previous study (Yu and Zhang, 2012).

Coding genes of denitrification enzymes include enzymes EC 1.7.1.1, EC 1.7.7.2 and EC 1.7.99.4, and their sequence

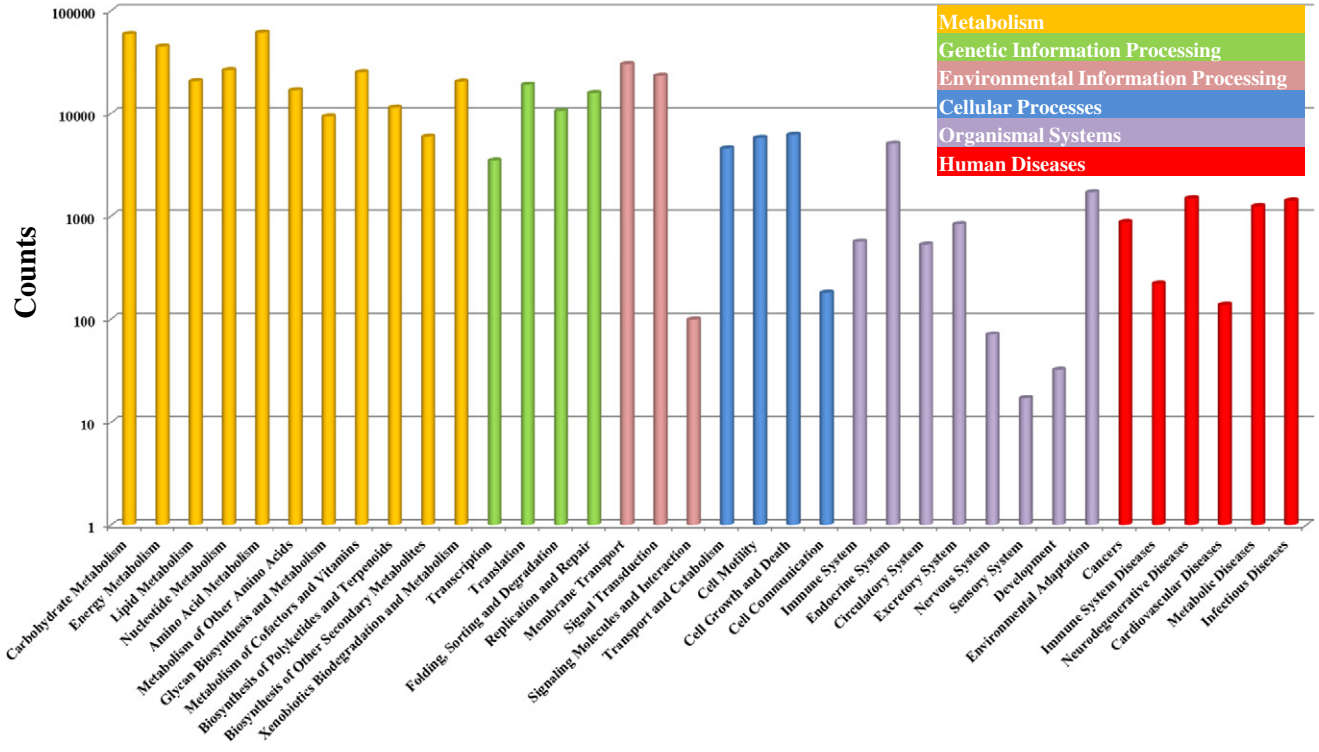

Fig. 5 – Functional classification and abundances from KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis in MEGAN (MEta Genome ANalyzer).

Table 1 – Top thirty most abundant genera in the A2O sludge metagenome.

Genus	Reads	Taxonomic position
Nitrospira	63658	Nitrospirae; Nitrospira <class>; Nitrospirales; Nitrospiraceae
Thauera	16882	Proteobacteria; Betaproteobacteria; Rhodocyclales; Rhodocyclaceae
Dechloromonas	15591	Proteobacteria; Betaproteobacteria; Rhodocyclales; Rhodocyclaceae
Ignavibacterium	10883	Ignavibacteria <phylum>; Ignavibacteria; Ignavibacteriales; Ignavibacteriaceae
Caldilinea	7616	Chloroflexi <phylum>; Caldilineae; Caldilineales; Caldilineaceae
Sorangium	7226	Proteobacteria; Deltaproteobacteria; Myxococcales; Sorangiineae; Polyangiaceae
Nitrosomonas	6585	Proteobacteria; Betaproteobacteria; Nitrosomonadales; Nitrosomonadaceae
Anaerolinea	5720	Chloroflexi <phylum>; Anaerolineae; Anaerolineales; Anaerolineaceae
Niastella	5493	Bacteroidetes; Sphingobacteriia; Sphingobacteriales; Chitinophagaceae
Haliscomenobacter	3944	Bacteroidetes; Sphingobacteriia; Sphingobacteriales; Saprospiraceae
Candidatus Accumulibacter	3656	Proteobacteria; Betaproteobacteria; unclassified Betaproteobacteria
Planctomyces	3467	Planctomycetes; Planctomycetia; Planctomycetales; Planctomycetaceae
Haliangium	2511	Proteobacteria; Deltaproteobacteria; Myxococcales; Nannocystineae; Kofleriaceae
Solibacter	2404	Acidobacteria; Solibacteres; Solibacterales; Solibacteraceae
Niabella	2162	Bacteroidetes; Sphingobacteriia; Sphingobacteriales Chitinophagaceae
Plesiocystis	2122	Proteobacteria; Deltaproteobacteria; Myxococcales; Nannocystineae Nannocystaceae
Burkholderia	1966	Proteobacteria; Betaproteobacteria; Burkholderiales; Burkholderiaceae
Rubrivivax	1935	Proteobacteria; Betaproteobacteria; Burkholderiales; unclassified Burkholderiales
Pedosphaera	1837	Verrucomicrobia; Verrucomicrobiae; Verrucomicrobiales; Verrucomicrobia subdivision 3
Hyphomicrobium	1800	Proteobacteria; Alphaproteobacteria; Rhizobiales; Hyphomicrobiaceae
Verrucomicrobium	1599	Verrucomicrobia; Verrucomicrobiae; Verrucomicrobiales; Verrucomicrobiaceae
Pirellula	1583	Planctomycetes; Planctomycetia; Planctomycetales; Planctomycetaceae
Acidovorax	1419	Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae
Gemmata	1340	Planctomycetes; Planctomycetia; Planctomycetales; Planctomycetaceae
Azoarcus	1322	Proteobacteria; Betaproteobacteria; Rhodocyclales; Rhodocyclaceae
Singulisphaera	1199	Planctomycetes; Planctomycetia; Planctomycetales; Planctomycetaceae
Pseudomonas	1109	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae
Aromatoleum	1083	Proteobacteria; Betaproteobacteria; Rhodocyclales; Rhodocyclaceae
Bdellovibrio	1057	Proteobacteria; Deltaproteobacteria; Bdellovibrionales; Bdellovibrionaceae
Chloracidobacterium	1036	Acidobacteria; unclassified Acidobacteria; Candidatus Chloracidobacterium

numbers are 31, 1 and 1540, respectively. There are 324 and 258 coding gene sequences of nitrous-oxide reductase (EC 1.7.2.4) and nitrite reductase (NO-forming) (EC:1.7.2.1), respectively. Ammonification enzymes coding genes of high sequence abundance include genes coding for nitrite reductase (EC 1.7.1.4, EC 1.7.7.1, and EC 1.7.2.2) (352 sequences) and hydroxylamine reductase (EC 1.7.99.1) (189 sequences). In nitrogen fixation process, 133 sequences are annotated as Nif (EC 1.18.6.1) coding gene. Of all the four processes, abundance of nitrification enzymes coding gene sequences is the lowest. There are only 58 and 66 sequences of genes coding for ammonia monooxygenase (EC 1.13.12.-) and Hao (EC 1.7.3.4), respectively. These results show that the genes related to nitrogen metabolism are widespread in the current A2O sludge.

Anaerobic ammonium oxidation (ANAMMOX) is a globally important microbial process of the nitrogen cycle (Arrigo, 2005). The bacteria mediating this process were identified in 1999, and were a great surprise for the scientific community at that time (Strous et al., 1999). In this biological process, nitrite and ammonium are converted into nitrogen gas directly. The bacteria that perform the ANAMMOX process belong to the bacterial phylum Planctomycetes. A total of 10,127 reads were detected in the phylum Planctomycetes, which belong to 11 genera. Among them, *Planctomyces* is the most abundant genus (3467 reads, 1.30%), followed by *Pirellula* (1583 reads, 0.60%), *Gemmata* (1340 reads, 0.50%) and *Singulisphaera* (1199 reads, 0.45%). All the four genera are ranked within the top thirty abundant genera (Table 1 & Fig. 2), revealing the important role of ANAMMOX process in nitrogen metabolism of the A2O sludge.

3. Conclusions

This is the first report based on Illumina sequencing analysis on the broad phylogenetic and metabolic diversity of A2O sludge metagenome from municipal sewage treatment plant. In this A2O active sludge ecosystem, 51 phyla and nearly 900 genera were identified, with considerable amount of high-quality unassigned sequences, which demonstrate the astonishing high biological diversity present in the A2O sludge. Proteobacteria, Bacteroidetes, Nitrospirae and Chloroflexi are predominant phyla in the active sludge. *Nitrospira*, *Thauera*, *Dechloromonas* and *Ignavibacterium* with the abilities to metabolize nitrogen and aromatic compounds are the most prevalent genera. Denitrification related genes have the highest sequence number, followed by ammonification, nitrogen fixation and nitrification, which revealed that denitrification coding gene sequences are dominant among the four processes of nitrogen metabolism.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 41430643).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jes.2014.12.027.

REFERENCES

Abu-Alhail, S., Lu, X.W., 2014. Experimental investigation and modeling of innovative five-tank anaerobic-anoxic/oxic process. Appl. Math. Model. 38 (1), 278–290.

Arrigo, K.R., 2005. Marine microorganisms and global nutrient cycles. Nature 437 (7057), 349–355.

Bae, H., Park, K.S., Chung, Y.C., Jung, J.Y., 2010. Distribution of anammox bacteria in domestic WWTPs and their enrichments evaluated by real-time quantitative PCR. Process Biochem. 45 (3), 323–334.

Baek, G., Cheon, S.P., Kim, S., Kim, Y., Kim, H., Kim, C., Kim, S., 2012. Modular neural networks prediction model based A(2)/O process control system. Int. J. Precis. Eng. Manuf. 13 (6), 905–913.

Brussow, H., Desiere, F., 2001. Comparative phage genomics and the evolution of Siphoviridae: insights from dairy phages. Mol. Microbiol. 39 (2), 213–222.

Cantalupo, P.G., Calgua, B., Zhao, G., Hundesa, A., Wier, A.D., Katz, J.P., Grabe, M., Hendrix, R.W., Girones, R., Wang, D., Pipas, J.M., 2011. Raw sewage harbors diverse viral populations. mBio 2 (5) (e00180-00111).

Chiellini, C., Munz, G., Petroni, G., Lubello, C., Mori, G., Verni, F., Vannini, C., 2013. Characterization and comparison of bacterial communities selected in conventional activated sludge and membrane bioreactor pilot plants: a focus on *Nitrospira* and *Planctomycetes* bacterial phyla. Curr. Microbiol. 67 (1), 77–90.

Coates, J.D., Chakraborty, R., Lack, J.G., O'Connor, S.M., Cole, K.A., Bender, K.S., Achenbach, L.A., 2001. Anaerobic benzene oxidation coupled to nitrate reduction in pure culture by two strains of *Dechloromonas*. Nature 411 (6841), 1039–1043.

DeLong, E.F., 2009. The microbial ocean from genomes to biomes. Nature 459 (7244), 200–206.

Fierer, N., Ladau, J., Clemente, J.C., Leff, J.W., Owens, S.M., Pollard, K.S., Knight, R., Gilbert, J.A., McCulley, R.L., 2013. Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. Science 342 (6158), 621–624.

Gilbert, E.M., Agrawal, S., Brunner, F., Schwartz, T., Horn, H., Lackner, S., 2014. Response of different nitrospira species to anoxic periods depends on operational do. Environ. Sci. Technol. 48 (5), 2934–2941.

Guan, X.Y., Wang, J.F., Zhao, H., Wang, J.J., Luo, X.M., Liu, F., Zhao, F.Q., 2013. Soil bacterial communities shaped by geochemical factors and land use in a less-explored area, Tibetan Plateau. BMC Genomics 14, 820.

Huson, D.H., Mitra, S., Ruscheweyh, H.J., Weber, N., Schuster, S.C., 2011. Integrative analysis of environmental sequences using MEGAN4. Genome Res. 21 (9), 1552–1560.

Jogler, C., Wanner, G., Kolinko, S., Niebler, M., Amann, R., Petersen, N., Kube, M., Reinhardt, R., Schuler, D., 2011. Conservation of proteobacterial magnetosome genes and structures in an uncultivated member of the deep-branching Nitrospira phylum. Proc. Natl. Acad. Sci. U. S. A. 108 (3), 1134–1139.

Juma, D., Wang, H.T., Li, F.T., 2014. Impacts of population growth and economic development on water quality of a lake: case study of Lake Victoria Kenya water. Environ. Sci. Pollut. Res. 21 (8), 5737–5746.

Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergstrom, G., Behre, C.J., Fagerberg, B., Nielsen, J., Backhed, F., 2013. Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature 498 (7452), 99–103.

Kim, B.C., Kim, S., Shin, T., Kim, H., Sang, B.I., 2013. Comparison of the bacterial communities in anaerobic, anoxic, and oxic chambers of a pilot A(2)O process using pyrosequencing analysis. Curr. Microbiol. 66 (6), 555–565.

Kostan, J., Sjoblom, B., Maixner, F., Mlynek, G., Furtmuller, P.G., Obinger, C., Wagner, M., Daims, H., Djinovic-Carugo, K., 2010. Structural and functional characterisation of the chlorite dismutase from the nitrite-oxidizing bacterium "*Candidatus* Nitrospira defluvii": Identification of a catalytically important amino acid residue. J. Struct. Biol. 172 (3), 331–342.

Kroeze, C., Hofstra, N., Ivens, W., Lohr, A., Strokal, M., van Wijnen, J., 2013. The links between global carbon, water and nutrient cycles in an urbanizing world — the case of coastal eutrophication. Curr. Opin. Environ. Sustain. 5 (6), 566–572.

Liaw, R.B., Cheng, M.P., Wu, M.C., Lee, C.Y., 2010. Use of metagenomic approaches to isolate lipolytic genes from activated sludge. Bioresour. Technol. 101 (21), 8323–8329.

Liu, Z.F., Frigaard, N.U., Vogl, K., Iino, T., Ohkuma, M., Overmann, J., Bryant, D.A., 2012. Complete genome of *Ignavibacterium album*, a metabolically versatile, flagellated, facultative anaerobe from the phylum Chlorobi. Front. Microbiol. 3, 185.

Liu, B.B., Mao, Y.J., Bergaust, L., Bakken, L.R., Frostegard, A., 2013. Strains in the genus *Thauera* exhibit remarkably different denitrification regulatory phenotypes. Environ. Microbiol. 15 (10), 2816–2828.

Mackelprang, R., Waldrop, M.P., DeAngelis, K.M., David, M.M., Chavarria, K.L., Blazewicz, S.J., Rubin, E.M., Jansson, J.K., 2011. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. Nature 480 (7377), 368–371.

Maixner, F., Wagner, M., Lucker, S., Pelletier, E., Schmitz-Esser, S., Hace, K., Spieck, E., Konrat, R., Le Paslier, D., Daims, H., 2008. Environmental genomics reveals a functional chlorite dismutase in the nitrite-oxidizing bacterium '*Candidatus* Nitrospira defluvii'. Environ. Microbiol. 10 (11), 3043–3056.

Mao, Y.J., Zhang, X.J., Yan, X., Liu, B.B., Zhao, L.P., 2008. Development of group-specific PCR-DGGE fingerprinting for monitoring structural changes of *Thauera* spp. in an industrial wastewater treatment plant responding to operational perturbations. J. Microbiol. Methods 75 (2), 231–236.

Mao, Y.P., Xia, Y., Zhang, T., 2013. Characterization of *Thauera*-dominated hydrogen-oxidizing autotrophic denitrifying microbial communities by using high-throughput sequencing. Bioresour. Technol. 128, 703–710.

Martin, H.G., Ivanova, N., Kunin, V., Warnecke, F., Barry, K.W., McHardy, A.C., Yeates, C., He, S.M., Salamov, A.A., Szeto, E., Dalin, E., Putnam, N.H., Shapiro, H.J., Pangilinan, J.L., Rigoutsos, I., Kyrpides, N.C., Blackall, L.L., McMahon, K.D., Hugenholtz, P., 2006. Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. Nat. Biotechnol. 24 (10), 1263–1269.

Metcalf, W.W., Griffin, B.M., Cicchillo, R.M., Gao, J.T., Janga, S.C., Cooke, H.A., Circello, B.T., Evans, B.S., Martens-Habbena, W., Stahl, D.A., van der Donk, W.A., 2012. Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. Science 337 (6098), 1104–1107.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., Edwards, R.A., 2008. The metagenomics RAST server — a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinf. 9, 386.

Mitra, S., Rupek, P., Richter, D.C., Urich, T., Gilbert, J.A., Meyer, F., Wilke, A., Huson, D.H., 2011. Functional analysis of metagenomes and metatranscriptomes using SEED and KEGG. BMC Bioinf. 12 (S1), S21.

Qin, J.J., Li, R.Q., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J.H., Xu, J.M., Li, S.C., Li, D.F., Cao, J.J., Wang, B.,

Liang, H.Q., Zheng, H.S., Xie, Y.L., Tap, J., Lepage, P., Bertalan, M., Batto, J.M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H.M., Yu, C., Li, S.T., Jian, M., Zhou, Y., Li, Y.R., Zhang, X.Q., Li, S.G., Qin, N., Yang, H.M., Wang, J., Brunak, S., Dore, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Bork, P., Ehrlich, S.D., Wang, J., 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464 (7285), 59–65.

Ridaura, V.K., Faith, J.J., Rey, F.E., Cheng, J., Duncan, A.E., Kau, A.L., Griffin, N.W., Lombard, V., Henrissat, B., Bain, J.R., Muehlbauer, M.J., Ilkayeva, O., Semenkovich, C.F., Funai, K., Hayashi, D.K., Lyle, B.J., Martini, M.C., Ursell, L.K., Clemente, J.C., Van Treuren, W., Walters, W.A., Knight, R., Newgard, C.B., Heath, A.C., Gordon, J.I., 2013. Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science 341 (6150), 1241214.

Salinero, K.K., Keller, K., Feil, W.S., Feil, H., Trong, S., Di Bartolo, G., Lapidus, A., 2009. Metabolic analysis of the soil microbe *Dechloromonas aromatica* str. RCB: indications of a surprisingly complex life-style and cryptic anaerobic pathways for aromatic degradation. BMC Genomics 10, 351.

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Mende, D.R., Kultima, J.R., Martin, J., Kota, K., Sunyaev, S.R., Weinstock, G.M., Bork, P., 2013. Genomic variation landscape of the human gut microbiome. Nature 493 (7430), 45–50.

Silva, C.C., Hayden, H., Sawbridge, T., Mele, P., Kruger, R.H., Rodrigues, M.V., Costa, G.G., Vidal, R.O., Sousa, M.P., Torres, A.P., Santiago, V.M., Oliveira, V.M., 2012. Phylogenetic and functional diversity of metagenomic libraries of phenol degrading sludge from petroleum refinery wastewater treatment system. AMB Express 2 (1), 18.

Strous, M., Fuerst, J.A., Kramer, E.H.M., Logemann, S., Muyzer, G., van de Pas-Schoonen, K.T., Webb, R., Kuenen, J.G., Jetten, M.S.M., 1999. Missing lithotroph identified as new planctomycete. Nature 400 (6743), 446–449.

Sun, H.W., Bai, Y., Peng, Y.Z., Xie, H.G., Shi, X.N., 2013. Achieving nitrogen removal via nitrite pathway from urban landfill leachate using the synergetic inhibition of free ammonia and free nitrous acid on nitrifying bacteria activity. Water Sci. Technol. 68 (9), 2035–2041.

Ushiki, N., Fujitani, H., Aoi, Y., Tsuneda, S., 2013. Isolation of *Nitrospira* belonging to sublineage II from a wastewater treatment plant. Microbes Environ. 28 (3), 346–353.

Walsh, D.A., Zaikova, E., Howes, C.G., Song, Y.C., Wright, J.J., Tringe, S.G., Tortell, P.D., Hallam, S.J., 2009. Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. Science 326 (5952), 578–582.

Wang, J.F., Qi, J., Zhao, H., He, S., Zhang, Y.F., Wei, S.C., Zhao, F.Q., 2013. Metagenomic sequencing reveals microbiota and its functional potential associated with periodontal disease. Sci. Rep. 3, 1843.

Winkler, M.K.H., Kleerebezem, R., de Bruin, L.M.M., Verheijen, P.J.T., Abbas, B., Habermacher, J., van Loosdrecht, M.C.M., 2013. Microbial diversity differences within aerobic granular sludge and activated sludge flocs. Appl. Microbiol. Biotechnol. 97 (16), 7447–7458.

Wu, C.Y., Peng, Y.Z., Wan, C.L., Wang, S.Y., 2011. Performance and microbial population variation in a plug-flow A(2)O process treating domestic wastewater with low C/N ratio. J. Chem. Technol. Biotechnol. 86 (3), 461–467.

Yang, Y., Yu, K., Xia, Y., Lau, F.T., Tang, D.T., Fung, W.C., Fang, H.H., Zhang, T., 2014. Metagenomic analysis of sludge from full-scale anaerobic digesters operated in municipal wastewater treatment plants. Appl. Microbiol. Biotechnol. 98 (12), 5709–5718.

Yu, K., Zhang, T., 2012. Metagenomic and metatranscriptomic analysis of microbial community structure and gene expression of activated sludge. PLoS One 7 (5), e38183.

Zhang, C.Q., Liang, Z.H., Hu, Z.Q., 2014. Bacterial response to a continuous long-term exposure of silver nanoparticles at sub-ppm silver concentrations in a membrane bioreactor activated sludge system. Water Res. 50, 350–358.