

Exploited Application of Pyrosequencing in Microbial Diversity of Activated Sludge System of Common Effluent Treatment Plants

M. P. Shah*

Industrial Waste Water Research Laboratory Division of Applied & Environmental Microbiology Enviro Technology Limited Gujarat, India

*Corresponding author: shahmp@uniphos.com

Received September 14, 2014; Revised October 11, 2014; Accepted October 13, 2014

Abstract Microbial Communities are actively present in the Activated Sludge System. We have applied PCR-based Pyrosequencing to investigate the bacterial communities of Activated Sludge samples from different common effluent treatment plants. A total of 259K effective sequences of 16S rRNA gene V4 region were obtained from these Activated Sludge samples. These sequences revealed huge amount of operational taxonomic units (OTUs) in Activated Sludge, that is, 1183–3567 OTUs in a sludge sample, at 3% cutoff level and sequencing depth of 16 489 sequences. Clear geographical differences among the Activated Sludge samples from effluent treatment Plant No.1 and No.2 were revealed by (1) cluster analyses based on abundances of OTUs or the genus/family/order assigned by Ribosomal Database Project (RDP) and (2) the principal coordinate analyses based on OTUs abundances, RDP taxa abundances and UniFrac of OTUs and their distances. In addition to certain unique bacterial populations in each Activated Sludge sample, some genera were dominant, and core populations shared by multiple samples, including two commonly reported genera of Zoogloea and Dechloromonas, three genera not frequently reported and three genera not well described so far. Pyrosequencing analyses of multiple Activated Sludge samples in this study also revealed the minority populations that are hard to be explored by traditional molecular methods and showed that a large proportion of sequences could not be assigned to taxonomic affiliations even at the phylum/class levels.

Keywords: pyrosequencing, activated sludge, bacteria, cluster analysis, PCoA

Cite This Article: M. P. Shah, “Exploited Application of Pyrosequencing in Microbial Diversity of Activated Sludge System of Common Effluent Treatment Plants.” *American Journal of Microbiological Research*, vol. 2, no. 5 (2014): 157-165. doi: 10.12691/ajmr-2-5-6.

1. Introduction

The growth of the world population, the development of various industries, and the use of fertilizers and pesticides in modern agriculture have overloaded not only the water resources but also the atmosphere and the soil with pollutants [Maulin et al., 2013]. Biological wastewater treatment is a multi-billion dollar industry—the largest application of biotechnology in the world. By harnessing and concentrating microorganisms in bioreactors, the beneficial activities of naturally-occurring microorganisms are accelerated, enabling removal of oxygen-depleting organics, toxics and nutrients, and preventing discharge of pathogens. Despite the environmental and economic importance of these processes, the knowledge of the microbial communities within biological wastewater treatment is incomplete, primarily because of a lack of suitable tools for their analysis. Culture-dependent methods are biased by the selection of species which obviously do not represent the real dominance structure [Gilbride et al, 2006; He Z et al, 2011]. The applications of conventional molecular

techniques in the last two decades, such as Fluorescence in situ Hybridization (FISH), Denaturing Gradient Gel Electrophoresis (DGGE), Terminal Restriction Fragment Length Polymorphism (T-RFLP), quantitative PCR, 16S rRNA clone libraries etc., have provided new insights into microbial community structure. However these methods generally are difficult to detect most of the low-abundance organisms in a wastewater microbial community, thus provide incomplete information about the microbial diversity and community within the biological WWTPs [Xia SQ et al, 2010]. Recently developed pyrosequencing technology can generate hundreds of thousands of short sequences [Margulies M et al., 2006], and significantly improve researchers’ ability to investigate the low abundance microorganisms [Sogin et al, 2006; McLellan, 2010; Park J, 2010; Peccia J, 2010; Qian PY, 2011]. This new approach has been extended to environmental systems in recent years [Fodor, 2014]. However, to our knowledge, there are rare reports using pyrosequencing technology to compare bacterial communities between various WWTPs operated at different geographic locations. More recently, Zhang et al. [15] and Xia et al. [7] have observed that some core populations of bacteria were shared by multiple activated sludge samples from different

wastewater treatment plants. However they did not examine the relationships between microbial community structures and environmental variables, thus it is not clear how wastewater characteristics, operational parameters and geographic locations influence the bacterial community structures. The low sequencing depth of the traditional PCR cloning approach when compared with the vast genetic diversity present in AS systems hindered a comprehensive characterization of the microbial community structure. In this aspect, the current community analyses typically represent a mere snapshot of the dominant members, with little information on taxa with medium to low abundances. High-throughput sequencing is a promising method, as it provides enough sequencing depth to cover the complex microbial communities [Shendure J et al, 2007]. So far, it has been applied to analyze microbial communities in marine water [Qian P, 2010], soil [Roesch L, 2007; Lauber CL, 2009], human hand surface [Fierer N, 2008], human distal intestine [Claesson M, 2009], etc. But few studies have been conducted using this method to investigate AS, though this approach has been used to study the microbial community in the raw sewage [McLellan S, 2010] and the residual biosolids after digestion of AS [Bibby K, 2010]. Using computational ecology tools such as the Uni Frac b-diversity metric [Lozupone CA, 2005] and principal coordinates analysis (PCoA), researchers can compare differences in microbial communities between ecosystems and along environmental gradients. With the aid of high-throughput sequencing technology and the well-established b-diversity analytical tool, an attempt could be made to answer some fundamental questions related to AS microbial communities. In this study, AS samples were collected from sewage treatment plants. Pyrosequencing using the 16S rRNA gene as the biomarker was conducted to examine the bacterial diversity of these AS samples, to evaluate the similarity/difference of different samples, to compare the unique dominant bacterial populations, and to identify the core populations shared by different samples. To the best of our knowledge, this study is the first application of PCR based pyrosequencing to characterize and compare multiple AS samples.

2. Materials & Methods

2.1 Waste Water Treatment Plant and Sampling

Biomass samples were collected from different effluent treatment plants of India. Grab samples were collected from the aerobic basin. A bucket-type sampler was used for the suspended microorganisms, whereas a sterilized scissor was used to cut off a part of the fixed synthetic mesh, after lifting the mesh, for the attached microorganisms. Samples were immediately frozen before being transported to the laboratory.

2.2. DNA Extraction

Of these samples, the same sludge sample from effluent treatment plant was divided into two aliquots, that is, A and B. Then the two subsamples were treated as independent samples from DNA extraction to pyrosequencing, to evaluate the reproducibility of the

methods applied in this study. Samples of 10ml were centrifuged at 4000 r.p.m. for 10 min at 4 °C. Pellet (200mg) of each sample was collected for DNA extraction in duplicate with the Fast DNA SPIN Kit for Soil, which was found to be the most suitable (having the lowest contamination) for the samples in this study. The duplicate DNA extracts were then merged together for the following PCR amplification.

2.3. PCR Amplification and Pyrosequencing

The forward primer 563F (50-AYTGGGYD TAAAGNG-30) at the 50-end (E. coli positions 563–578) of the V4 region (239 nucleotides) and a cocktail of four equally mixed reverse primers, that is, R1 (50-TACCRGGGTHCTAATCC-30), R2 (50-TAC CAGAGTATCTAATTC-30), R3 (50-CTACDSRGGTMTCTAATC-30) and R4 (50-TACNVTGGGTATCTAATC-30), at the 30-end of the V4 region (E. coli positions 785–802) were selected for PCR, because they may capture 95% of all bacterial 16S rRNA gene sequences in databases (Murphy et al., 2010). They have also been used in other microbial diversity studies employing the PCR-based Pyrosequencing method (Claesson et al., 2009; Jesus et al., 2010; Murphy et al., 2010). The 50-fused primer includes a 10 nucleotide 'barcode' inserted between the Roche 454 life science adapter primer A and the 563F primer. The barcode is permuted for each sample and allows the identification of individual samples in a mixture in a single Pyrosequencing run (Sogin et al., 2006). A 100 ml reaction system was set up for each PCR amplification, using Mighty Amp polymerase. The amplification was conducted in an i-Cycler under the following conditions: initial denaturation at 98°C for 2 min, and 28 cycles at 98°C for 15 s, 56°C for 20 s and 68°C for 30 s, and a final extension at 68°C for 10 min. Amplicon libraries were prepared by a cocktail of three independent PCR products for each sample to minimize the impact of potential early-round PCR errors (Sogin et al., 2006). PCR amplicons were purified with a quick-spin Kit, and concentrations were measured using spectrometry. Amplicons from different sludge samples were then mixed to achieve equal mass concentrations in the final mixture, which was sent out for pyrosequencing at Bangalore Genei, India.

2.4. Data Analysis

Initially, the raw sequences obtained from pyrosequencing were processed using the GL FLX software for sorting by the key (i.e., sequences from the suspended and attached samples), with low quality sequences discarded, and the primer sequences trimmed. In addition, sequences less than 250-bp long were discarded. The processed sequences were subjected to taxonomical classification

2.5. Post-run Analysis

All the raw reads were treated with the Pyrosequencing Pipeline Initial Process (Cole et al., 2009) of the Ribosomal Database Project (RDP), (1) to sort those exactly matching the specific barcodes into different samples, (2) to trim off the adapters, barcodes and primers using the default parameters, and (3) to remove sequences

containing ambiguous 'N' or shorter than 150 bps (Claesson et al., 2009). The reads selected above were defined as 'raw reads' for each AS sample. Sequences were denoised using the 'pre.cluster' command in Mothur platform to remove sequences that are likely due to pyrosequencing errors (Huse et al., 2010; Roeselers et al., 2011). PCR chimeras were filtered out using Chimera Slayer (Haas et al., 2011). The reads flagged as chimeras was extracted out and submitted to RDP. Those being assigned to any known genus with 90% confidence were merged with the non-chimera reads, to form the collection of 'effective reads' for each AS sample. Although bacteria-specific primers were applied, very small portions of unexpected archaeal sequences might be obtained (Qian et al., 2010). To remove these cross-talking sequences, the effective sequences of each AS sample were submitted to the RDP Classifier (Wang et al., 2007) to identify the archaeal and bacterial sequences, and filter out the archaeal sequences using a self-written Python script. The average length of all bacterial sequences without the primers was 207 bp. After the above filtration, the minimum number of selected bacterial sequences in the 15 AS samples was 16 489. To fairly compare the different Activated Sludge samples at the same sequencing depth, normalization of the sequence number was conducted by extracting the first 16 489 sequences in each sample for all the following analyses. Taxonomic classification of the bacterial sequences of each Activated Sludge sample was carried out individually, using the RDP Classifier. A bootstrap cut off of 50% suggested by the RDP was applied to assign the sequences to different taxonomy levels. The normalized sequence set of each AS sample was also individually aligned by Infernal (Nawrocki and Eddy, 2007) using the bacteria-alignment model in Align module of the RDP. By applying the Complete Linkage Clustering, sequences in each set were assigned to phylotype clusters at two cutoff levels of 3% and 6%. The cluster analysis (CA) was conducted to group the bacterial communities of different Activated Sludge samples on the basis of (1) taxonomy results obtained using the RDP Classifier (excluding those unclassified sequences), and (2) operational taxonomic units (OTUs) generated using RDP Complete Linkage Clustering from the merged pool of sequences of all the Activated Sludge samples. A sample mapping file showing the frequency of each reference taxon in different Activated Sludge samples was generated through local BLAST.

3. Results and Discussion

3.1. Microbial Community Diversity

As shown in Table 1, after filtering the low quality reads using the RDP Initial Process in Pyrosequencing Pipeline (PP) and trimming the adapters, barcodes and primers, there were 20 276B28 260 effective reads for the different Activated sludge samples. After denoising, filtering out chimeras, and removing the archaeal sequences, the library size of each sample was normalized to 16 489 sequences, which were the smallest among the different Activated sludge samples, to conduct the downstream analyses for different samples at the same sequencing depth. The numbers of OTUs, Chao 1 and

ACE at two cutoff levels of 3% and 6% are summarized in Table 1. On the basis of the OTU number, the Activated sludge sample from effluent treatment plant No.1 had the richest diversity, followed closely by those from No.2 and No.3, whereas the three samples from No.4, No.5 and No.6 displayed considerably less richness. The patterns of Chao 1 and ACE values were very similar to the OTU numbers. All three indices, that is, OTU number, Chao 1 and ACE, demonstrated that the richness values varied by 2–3 times among these Activated sludge samples. Plots of OTU number versus sequence number, that is, the rarefaction curves, are shown in Figure 2. As a common hypothesis, diversity may affect the performance of the Activated sludge process. However, an accurate estimation of OTUs in an Activated sludge sample, based on DNA sequencing, has not been conducted before. In this study, based on the 247 335 effective bacterial sequences, there could be a total of 13 951 (3% cutoff) and 8493 OTUs (6%) in the 15 samples. It should be noted that the richness values were certainly affected by sequencing noise. The 16 489 selected effective bacterial sequences in each samples were assigned to different taxa levels (from genus to phylum) using the RDP Classifier at 50% threshold. Although it has been reported that the V4 region used in this study displayed the highest number of correctly classified sequences, followed by the V3 and V6 regions (Claesson et al., 2009), quite a large portion of effective bacterial sequences in this study could not be assigned to any taxa of different level at the 50% threshold, indicating the extent of novel sequences captured by this study. Figure 1 shows that the unclassified sequence portions in the total community increased from the domain level to the genus level, and were significantly different among the samples, especially at the family and genus levels. As shown in Figure 1, Proteobacteria was the most abundant phylum in all samples, accounting for 36–65% of total effective bacterial sequences. This is similar to the analytical results of bacterial communities in soil (Roesch et al., 2007) and sewage (McLellan et al., 2010), in which Proteobacteria was also the most dominant community. The other dominant phyla were Firmicutes (1.4–14.6%, averaging at 8.1%), Bacteroidetes (2.7–15.6%, averaging at 7.0%) and Actinobacteria (1.3–14.0%, averaging at 6.5%). Similar to a few previous studies on Activated Sludge using microarray (Xia et al., 2010) and cloning (Snaird et al., 1997), these four groups were dominant (56–86%) in bacterial communities of the different Activated Sludge samples in this study, followed by a few other major (average abundance 41%) phyla, including Verrucomicrobia (4.2%), Chloroflexi (3.4%), Acidobacteria (3.0%) and Planctomycetes (2.4%). A few phyla, including TM7, Thermotogae, OD1, Spirochaetes, WS3, Nitrospira and Synergistetes, were only the major (abundance 41%) phyla in one of the different samples. The abundances of other phyla were 0.1% in all samples. Within Proteobacteria, e-Proteobacteria only occurred at very low levels (0.01–0.51%, averaging 0.08%). Except for the four samples of effluent treatment no.5, which had the a-subdivision as the most dominant class within Proteobacteria, in all other 11 samples, the b-subdivision was the most dominant Proteobacteria, followed by a-, g- and d-subdivisions. This is different from results of a study using microarray (Xia et al., 2010), which showed that a-subdivision was the most abundant of the

Proteobacteria. However, the findings are similar to the results of another study about soil bacteria using pyrosequencing (Roesch et al., 2007), which demonstrated that in most soils, the b-subdivision was the most abundant one within the Proteobacteria. It is interesting that in the Activated Sludge sample from the unique saline effluent treatment plant at Ankleshwar, India the a subdivision was the dominant class within the Proteobacteria, accounting for 18.5–25.7% of total bacterial effective sequences, whereas the a-Proteobacteria only averaged 10.7% in the other 13 Activated Sludge samples. This might be explained by the dominance (up to 25–50%) of the a-Proteobacteria in the marine microbial

community of the surface/subeuphotic layers (Venter et al., 2004; Qin et al., 2010). In addition, the d-Proteobacteria was the most dominant class in the other two effluent treatment plants of India, but their most dominant orders in the d-Proteobacteria were different, that is, Myxococcales (11%) and Desulfobacterales (5.8%), respectively. In addition to the four classes of the Proteobacteria, other dominant (41%) shared classes included Actinobacteria, Sphingobacteria, Clostridia, Bacilli, Planctomycetacia, Verrucomicrobiae, Anaerolineae, Verrucomicrobia Subdivision3, Flavobacteria, Caldilineae, Acidobacteria_Gp4 and Acidobacteria_Gp6.

Table 1. Percentages of the shared genera and their corresponding sequences

Number of sample11	Number of shared genera	Percentage in classified generah	Percentage in classified sequenced
15	70	9.5	63.7
14	121	16.4	81.2
13	158	21.4	86.5
12	184	24.9	89.3
11	208	28.1	91.0
10	235	31.8	92.7
9	253	34.2	93.6
8	278	37.6	94.0
7	305	41.3	94.5
6	339	45.9	97.3
5	373	50.5	97.7
4	404	54.7	98.4
3	468	63.3	99.1
2	556	75.2	99.7
1	739	100.0	100.0

Abbreviation: AS, activated sludge.

Among the 744 assigned genera, 70 (accounting for 63.7% of the classified sequences) were shared by all 15 samples. A total of 235 genera were commonly shared by more than 10 AS samples, accounting for 92.7% of all classified sequences. There were 273 rare genera that only appeared in one or two samples, accounting for only 0.9% of total sequences.

^aNumber of AS samples, which share the genera in the second column.

^bPercentage of the number of shared genera in the number of total classified genera.

^cPercentage of sequences of the shared genera in the total classified sequences.

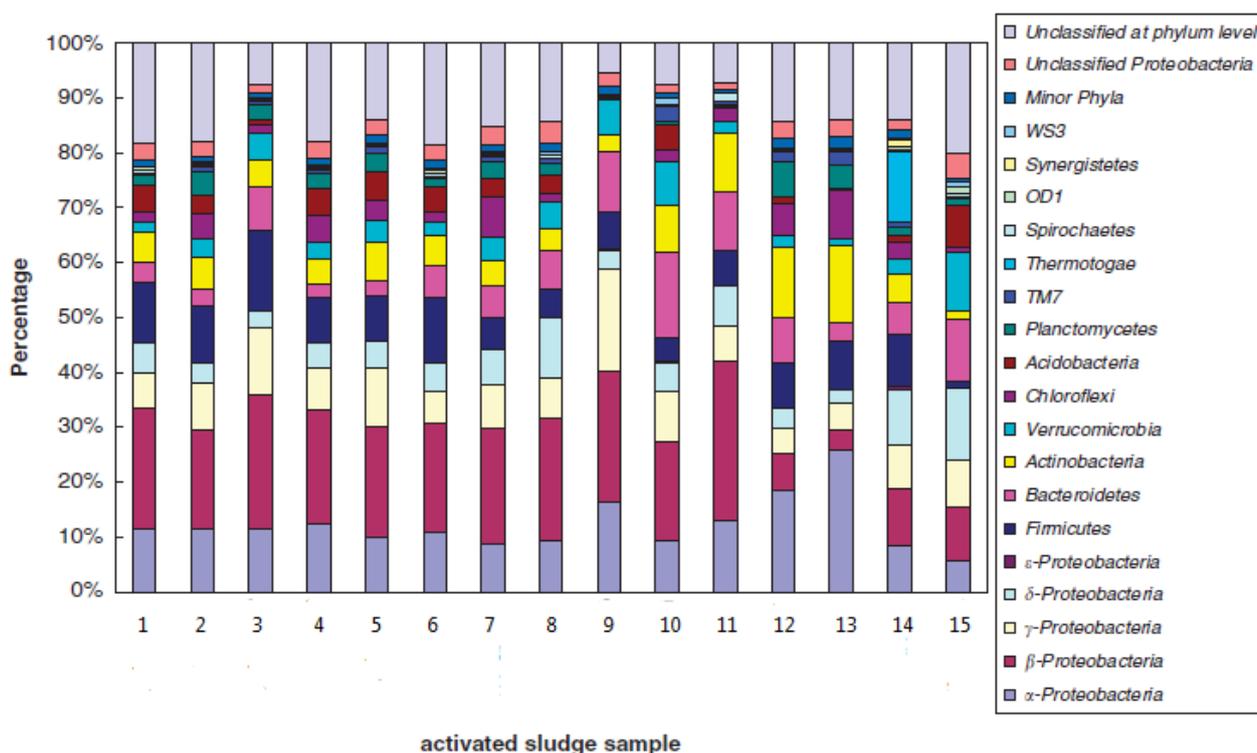


Figure 1. Abundances of different phyla and classes in Proteobacteria in the 15 Activated Sludge samples

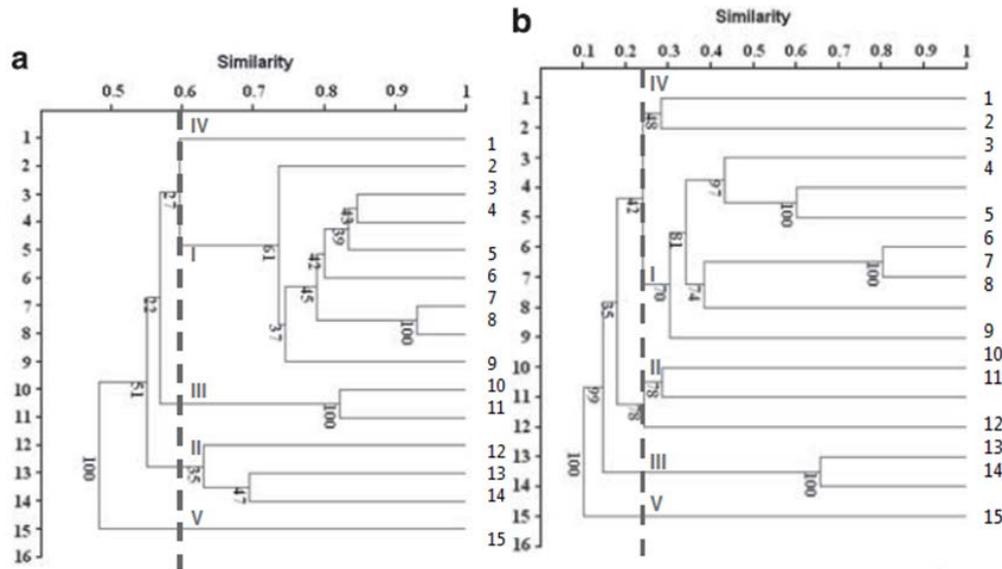


Figure 2. CA based on Bray-Curtis distances of 15 Activated Sludge samples. (a) At order level; (b) at 3% cut off-OTU level. The dot lines show the similarity cut off levels to cluster the 15 AS samples into five groups

3.2. Analysis of Similarity

The similarity of the Activated Sludge samples was evaluated using two independent methods: CA and PCoA. Cluster analysis. As shown in Figure 2a, CA, based on abundances of orders, revealed that bacterial communities in the 15 samples could be clustered into five groups: (1) Group I contains all Activated Sludge samples from effluent treatment plant No.1 and that from No.2; (2) Group II contains the three samples from No.3; (3) Group III is sludge from No.4, which treats saline sewage; (4) Group IV is sludge from No.5, which treats sewage containing slaughterhouse wastewater; (5) Group V is sludge from No.6, which is located inside a cave. This grouping pattern was similar at the family and the genus levels (Figure 4b, 4c), but slightly changed at the class levels (Figure 4a), and eventually disappeared at the phylum level. Using 0.6 as a benchmark (McLellan et al., 2010), the bacterial compositions (at the family/order levels) of the different sludge samples are quite similar in Group I, even though the ETPs in different parts of India are separated by over thousands km. Using the OTUs abundances, similar grouping patterns were observed at

both 3% (Figure 2b) and 6% cutoffs (Figure 4d). Principal coordinates analysis. PCoA was also conducted to evaluate similarities of different Activated Sludge samples using three different approaches, that is, RDP Classifier taxa, OTUs and UniFrac. For the first two approaches, taxa or OTUs are regarded as equally related, whereas UniFrac incorporates the degree of divergence in the phylogenetic tree of OTUs into PCoA (Hamady et al., 2010; Qian et al., 2010). The PCoA analysis results are shown in Figure 3 (UniFrac at 3% cut off) and Figure 5a (order) and Figure 5b (OTUs at 3% cutoff). PCoA at the family, genus and 6% cutoff levels were also conducted. Although there are slight variances among these PCoA results, the same general trend was observed, that is, the sludge samples from ETPs forming clusters different from the sludge from other ETPs. We hypothesized that there would be significant differences among the sludge from different geographical areas. As demonstrated by the CA and PCoA, the sludge samples were certainly similar to those from the same geographical area, possibly due to unique sewage compositions, as well as different plant design and operation in each area.

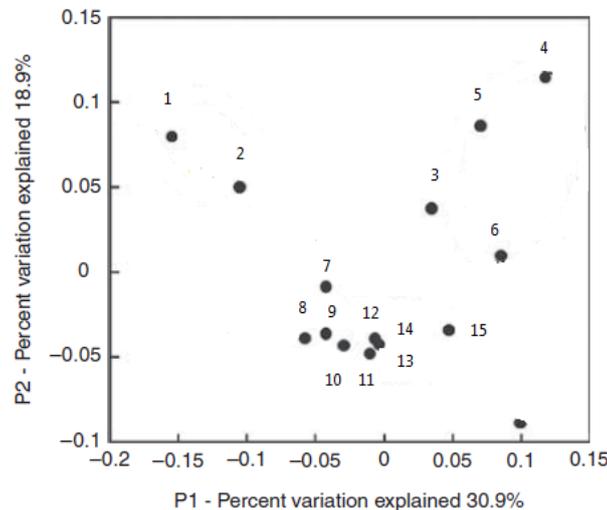


Figure 3. Principal coordinates analysis of 15 Activated Sludge samples by weighted UniFrac. For UniFrac, Green genes coresets tree was selected as the reference tree, a sample mapping file showing the frequency of each reference taxon in different AS samples was generated through local BLAST

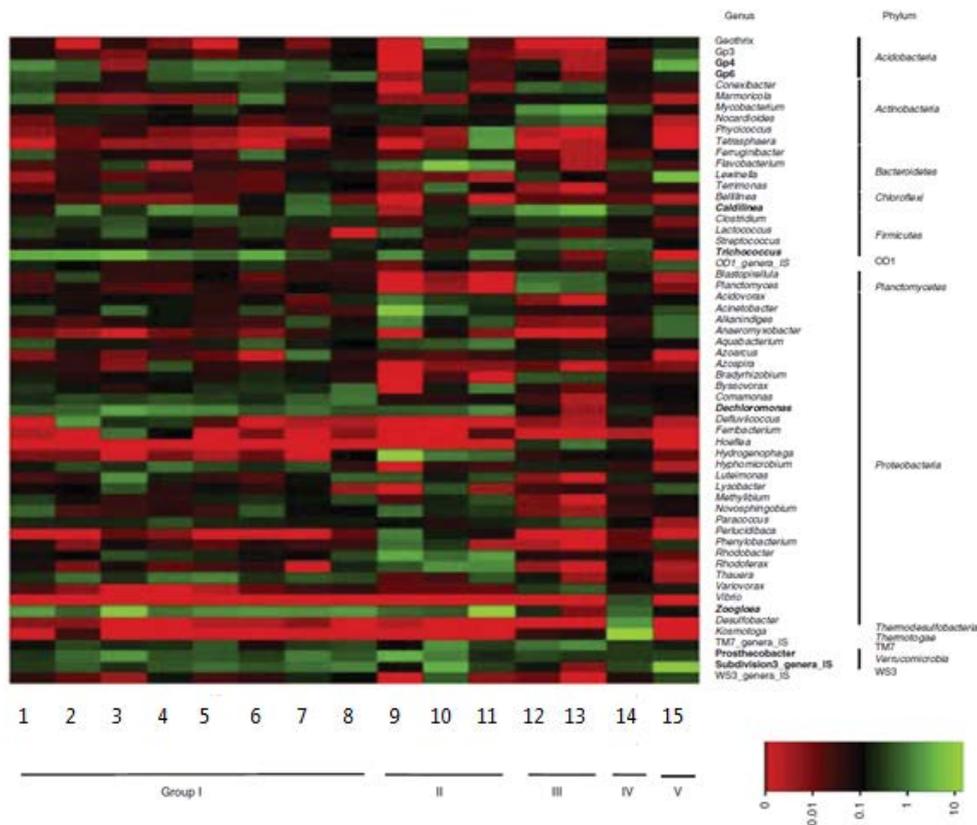


Figure 4. Heat map of top 10 genera in each sample. The top 10 abundant genera in each sample were selected (a total of 58 genera for all 15 samples) and compared with their abundances (percentages) in other samples. The color intensity (log scale) in each panel shows the percentage of a genus in a sample, referring to color key at the right bottom. Those in bold font are the core genera in different AS samples

3.3. Shared and Distinct Orders/Families

At the order level, the 27 abundant (41% in any Activated Sludge sample) orders accounted for 45–80% of the classified sequences, as shown in Figure 3. Among the 27 orders, Rhodocyclales, Burkholderiales, Rhizobiales, Myxococcales, Clostridiales, Sphingobacteriales, Actinomycetales, Rhodobacteriales, Xanthomonadales, Planctomycetales, Pseudomonadales and Verrucomicrobiales were the orders commonly shared by all sludges. However, two orders were significantly ($P < 0.1$) more abundant in effluent treatment plant no.1, that is, Burkholderiales and Sphingobacteriales, whereas three (i.e., Rhizobiales, Planctomycetales and Desulfobacteriales) were more abundant in sludge of effluent treatment plant no.2. CA after removing these distinct five orders shows mixed clusters of sludge from the two geographical areas. The CA based on these data showed distinct geographical clusters for Activated Sludge samples. Nocardioideaceae, Streptococcaceae, Cystobacteriaceae, Polyangiaceae, Rhodocyclaceae and Verrucomicrobiaceae were the families commonly shared by all sludge, whereas Flavobacteriaceae, Comamonadaceae and Sphingomonadaceae were significantly more abundant in no.1 sludge. Anaerolineaceae, Planctomycetaceae, Bradyrhizobiaceae, Desulfobacteraceae, Hyphomicrobiaceae and Rhodospirillaceae were more abundant in no.2 sludge.

3.4. Distinct Genera

Figure 5 shows the genus profiles for the fifteen Activated Sludge samples. Each panel represents one of

the 739 genera, which were sorted alphabetically by phyla and then genus. Examination of these genera captured the similarities and differences in the genus profiles of the 15 samples. Comparative analysis revealed a core microbiota across the 15 Activated Sludge samples. Among the 744 assigned genera, 70 (accounting for 63.7% of the classified sequences) were shared by all 15 samples. A total of 235 genera were commonly shared by more than 10 Activated Sludge samples, accounting for 92.7% of all classified sequences. There were 271 rare genera that only appeared in one or two samples, accounting for only 0.9% of total classified sequences, a very minor part of the bacterial communities in Activated Sludge. The top 10 abundant genera in each sample were selected (a total of 58 genera for all 15 samples) and compared with their abundances in other samples, as shown in Figure 4. Eight genera were abundant (41%) in at least six samples, including two genera extensively reported in Activated Sludge, that is, Zoogloea and Dechloromonas, three genera rarely reported before, that is, Caldilinea, Tricoccus and Prostheco bacter, plus three not well described genera, that is, Gp4, Gp6 and Subdivision3_genera_incertain_sedis. The species in the genus Zoogloea, such as Zoogloea ramigera, are known to form characteristic cell aggregates embedded in extracellular gelatinous matrices, often called zoogloal matrices (Dugan et al., 1992), and are the main agent for the flocculation of Activated Sludge (Rossello-Mora et al., 1995). Caldilinea has some filamentous species and has a role in stabilizing flocs of Activated Sludge in a wide range of STPs (Yoon et al., 2010). The information about the existence and roles of other genera in AS are limited. Just like at the family level, the dominant genera showed

some geographical characteristics. For example, Flavobacterium were only dominant in the three samples from Plant No.2, with abundance levels ranging from 1.83 to 7.44%, whereas its abundances were 0% in all the samples from Plant No.3 & 4. In addition, all three samples had high levels of the photosynthetic bacteria Rhodobacter (1.43–3.72%), whereas the samples from plant no. 3 & 4 contained relatively less (0.32–0.99%). As shown in Figure 5, a few other genera more dominant included Acinetobacter, Rhodofera, Acidovorax, Aquabacterium and Alkanindiges, whereas Gp4, Gp6, Clostridium, Hyphomicrobium, Comamonas and Pirellula were more abundant. The distribution of some abundant genera also depended on temperature. For instance, the species in Trichococcus genus have been identified as psychrotolerant mesophiles and are able to grow at low temperatures. The sludge samples also contained some other unique populations, possibly due to salinity,

exposure to sunlight, industrial wastewater contribution and so on. Two of the samples from different effluent treatment plant, which had been exposed to seawater, contained unique genera such as Hoeflea, of which all three known species are related to marine environments (Palacios et al., 2006). The unique populations in different Activated Sludge samples could be due to many possible causes, including geographical isolation, concentration and types of organic substrates in the sewage, which are significantly affected by the diets of eastern and western peoples, operation mode (A/A/O or A/O vs conventional AS), dissolved oxygen concentration, temperature, salinity, season, type and percentage of industrial wastewater, etc. Although it is out of the scope of the present study, the factors shaping microbial community deserve more comprehensive and systematic studies in the future, using the methodology demonstrated here.

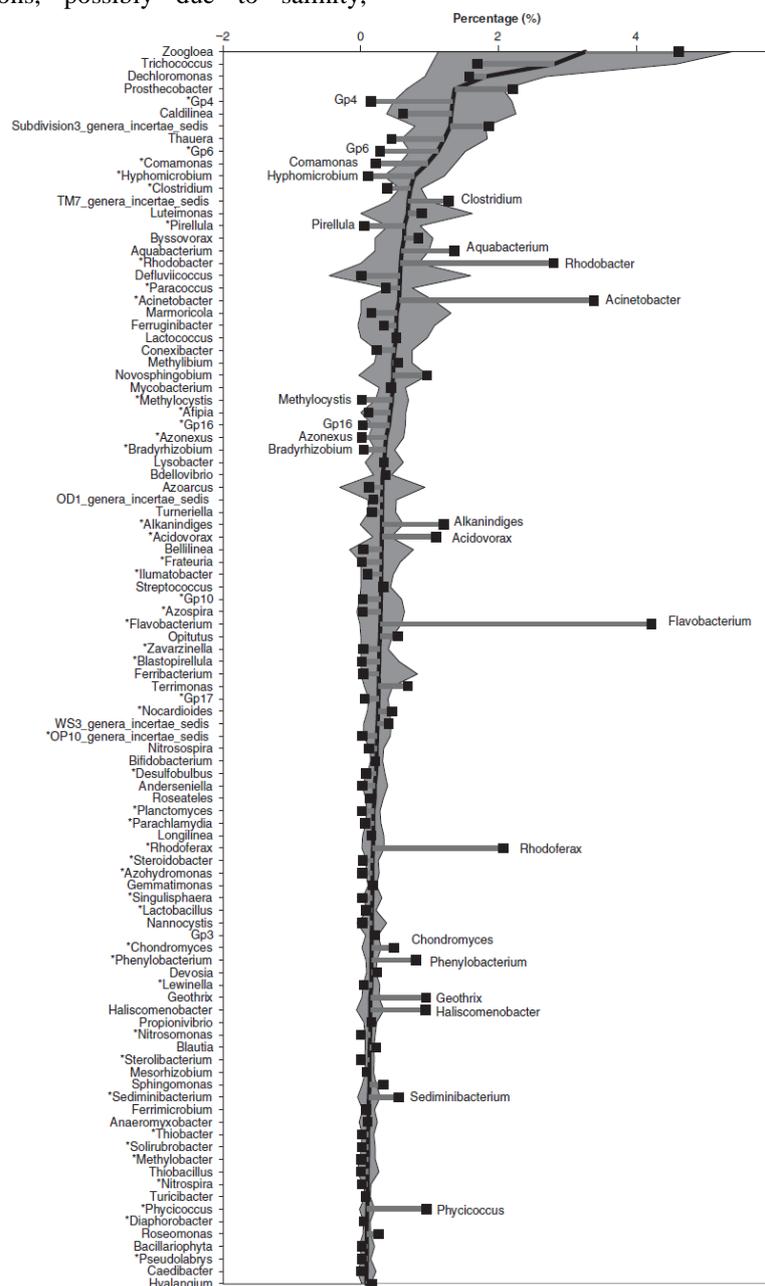


Figure 5. Average abundances (percentages) of the top 100 genera in eight AS samples of Group I, and their average percentages in three AS samples of Group II. The blue line shows the average percentages in Group I and the shadow area shows the variation ranges (average \pm SD). The up-bars and down-bars show the average percentages of the corresponding genera in Group II. *Shows the genera, which had significantly different average abundances in Groups I and II

References

- [1] Bibby K, Viau E, Peccia J. (2010). Pyrosequencing of the 16S rRNA gene to reveal bacterial pathogen diversity in biosolids. *Water Res* 44: 4252-4260.
- [2] Claesson M, O'Sullivan O, Wang Q, Nikkila J, Marchesi J, Smidt H et al. (2009). Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One* 4: e6669.
- [3] Claesson M, O'Sullivan O, Wang Q, Nikkila J, Marchesi J, Smidt H et al. (2009). Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One* 4: e6669.
- [4] Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ et al. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37: D141-D145.
- [5] Dugan PR, Stoner DL, Pickrum HM. (1992). The genus *Zoogloea*. In: *The prokaryotes*. Balows A, Tru" per HG, Dworkin M, Harder W, Schleifer K-H (eds). Springer-Verlag: New York, NY, pp 3952-3964.
- [6] Fierer N, Hamady M, Lauber CL, Knight R. (2008). The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci USA* 105: 17994.
- [7] Fodor AA, Sanapareddy N, Hamp TJ, Gonzalez LC, Hilger HA, and Clinton SM. 2009. Molecular Diversity of a North Carolina Wastewater Treatment Plant as Revealed by Pyrosequencing. *Appl. Environ. Microb.* 75: 1688-1696.
- [8] Gilbride KA, Lee DY, and Beaudette LA. 2006. Molecular techniques in wastewater: Understanding microbial communities, detecting pathogens, and real-time process control. *J. Microbiol. Meth.* 66: 1-20.
- [9] Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G et al. (2011). Chimeric 16S Rrna sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 21: 494-504.
- [10] Hamady M, Lozupone C, Knight R. (2010). Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of Pyrosequencing and PhyloChip data. *ISME J* 4: 17-27.
- [11] He Z, Van Nostrand JD, Deng Y, and Zhou J. 2011. Development and applications of functional gene microarrays in the analysis of the functional diversity, composition, and structure of microbial communities. *Front. Environ. Sci. Engin. China* 5:1-20.
- [12] Huse SM, Welch DM, Morrison HG, Sogin ML. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* 12: 1889-1898.
- [13] Jesus EC, Susilawati E, Smith S, Wang Q, Chai B, Farris R et al. (2010). Bacterial communities in the rhizosphere of biofuel crops grown on marginal lands as evaluated by 16S rRNA gene pyrosequences. *BioEnergy Res* 3:20-27.
- [14] Lauber CL, Hamady M, Knight R, Fierer N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl Environ Microbiol* 75: 5111.
- [15] Lozupone CA, Knight R. (2005). Unifrac: A new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71: 8228-8235.
- [16] Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen ZT, Dewell SB, de Winter A, Drake J, Du L, Fierro JM, Forte R, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Hutchison SK, Izzyk GP, Jando SC, Alenquer MLL, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lee WL, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Reifler M, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Willoughby DA, Yu PG, Begley RF, and Rothberg JM. 2006. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 441:120-120.
- [17] Maulin P Shah, Patel KA, Nair SS, Darji AM, Shaktisinh Maharaul. Optimization of Environmental Parameters on Decolorization of Remazol Black B Using Mixed Culture. *American Journal of Microbiological Research*. 2013 (1), 3, 53-56
- [18] Maulin P Shah, Patel KA, Nair SS, Darji AM. Microbial Decolorization of Methyl Orange Dye by *Pseudomonas spp.* *ETL-M. International Journal of Environmental Bioremediation and Biodegradation*. 2013 (1), 2, 54-59
- [19] Maulin P Shah, Patel KA, Nair SS, Darji AM. Microbial Degradation and Decolorization of Reactive Orange Dye by Strain of *Pseudomonas Spp.* *International Journal of Environmental Bioremediation and Biodegradation*. 2013 (1), 1, 1-5
- [20] Maulin P Shah, Patel KA, Nair SS, Darji AM. An Innovative Approach to Biodegradation of Textile Dye (Remazol Black) by *Bacillus spp.* *International Journal of Environmental Bioremediation and Biodegradation*. 2013 (1), 2, 43-48
- [21] McLellan S, Huse S, Mueller Spitz S, Andreishcheva E, Sogin M. (2010). Diversity and population structure of sewage derived microorganisms in wastewater treatment plant influent. *Environ Microbiol* 12: 378-392.
- [22] McLellan SL, Huse SM, Mueller-Spitz SR, Andreishcheva EN, and Sogin ML. 2010. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environ. Microbiol.* 12: 1376-1376.
- [23] Murphy E, Cotter P, Healy S, Marques T, O'Sullivan O, Fouhy F et al. (2010). Composition and energy harvesting capacity of the gut microbiota: relationship to diet, obesity and time in mouse models. *Gut* 59:1635-1642.
- [24] Nawrocki EP, Eddy SR. (2007). Query-dependent banding (QDB) for faster RNA similarity searches. *PLoS Comput Biol* 3: e56.
- [25] Palacios L, Arahal D, Reguera B, Marin I. (2006). *Hoeflea alexandrii* sp. nov., isolated from the toxic dinoflagellate *Alexandrium minutum* AL1V. *Int J Syst Evol Microbiol* 56: 1991.
- [26] Park J, Lee TK, Doan TV, Yoo K, Choi S, and Kim C. 2010. Discovery of commonly existing anode biofilm microbes in two different wastewater treatment MFCs using FLX Titanium pyrosequencing. *Appl. Microbiol. Biot.* 87: 2335-2343.
- [27] Peccia J, Bibby K, and Viau E. 2010. Pyrosequencing of the 16S rRNA gene to reveal bacterial pathogen diversity in biosolids. *Water Res.* 44: 4252-4260.
- [28] Qian P, Wang Y, Lee O, Lau S, Yang J, Lafi F et al. (2010). Vertical stratification of microbial communities in the Red Sea revealed by 16S rDNA pyrosequencing. *ISME J* 5: 507-518.
- [29] Qian P, Wang Y, Lee O, Lau S, Yang J, Lafi F et al. (2010). Vertical stratification of microbial communities in the Red Sea revealed by 16S rDNA pyrosequencing. *ISME J* 5: 507-518.
- [30] Qian PY, Wang Y, Lee OO, Lau SCK, Yang JK, Lafi FF, Al-Suwailem A, and \ Wong TYH. 2011. Vertical stratification of microbial communities in the Red Sea revealed by 16S rDNA pyrosequencing. *ISME. J.* 5: 507-518.
- [31] Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464: 59-65.
- [32] Roesch L, Fulthorpe R, Riva A, Casella G, Hadwin A, Kent A et al. (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* 1: 283-290.
- [33] Roeselers G, Mitghe EK, Stephens WZ, Parichy DM, Cavanaugh CM, Guillemin K et al. (2011). Evidence for a core gut microbiota in the zebrafish. *ISME J* 5: 1595-1608.
- [34] Rossello-Mora R, Wagner M, Amann R, Schleifer K. (1995). The abundance of *Zoogloea ramigera* in sewage treatment plants. *Appl Environ Microbiol* 61: 702.
- [35] Shendure J, Ji H. (2008). Next-generation DNA sequencing. *Nat Biotechnol* 26: 1135-1145.
- [36] Snaird J, Amann R, Huber I, Ludwig W, Schleifer KH. (1997). Phylogenetic analysis and in situ identification of bacteria in activated sludge. *Appl Environ Microbiol* 63: 2884-2896.
- [37] Sogin M, Morrison H, Huber J, Welch D, Huse S, Neal P et al. (2006). Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Nat Acad Sci* 103: 12115.
- [38] Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, Arrieta JM, and Herndl GJ. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. USA.* 103: 12115-12120.
- [39] Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304: 66-74.
- [40] Wang Q, Garrity G, Tiedje J, Cole J. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73: 5261-5267.
- [41] Xia S, Duan L, Song Y, Li J, Piceno Y, Andersen G et al. (2010). Bacterial community structure in geographically distributed

- biological wastewater treatment reactors. *Environ Sci Technol* 44: 1043-1045.
- [42] Xia SQ, Duan LA, Song YH, Li JX, Piceno YM, Andersen GL, Alvarez-Cohen L, Moreno-Andrade I, Huang CL, and Hermanowicz SW. 2010. Bacterial Community Structure in Geographically Distributed Biological Wastewater Treatment Reactors. *Environ. Sci. Technol.* 44: 7391-7396.
- [43] Yoon DN, Park SJ, Kim SJ, Jeon CO, Chae JC, Rhee SK. (2010). Isolation, characterization, and abundance of filamentous members of Caldilineae in activated sludge. *J Microbiol* 48: 275-283.
- [44] Zhang T, Shao MF, and Ye L. 2012. 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *ISME J.* 6: 1137-1147.