Environ. Sci. Technol. **2010,** *44,* 7391–7396

Bacterial Community Structure in Geographically Distributed Biological Wastewater Treatment Reactors

SIQING XIA,[*,†] LIANG DUAN,[‡] YONGHUI SONG,[‡] JIXIANG LI,[†] YVETTE M. PICENO,[||] GARY L. ANDERSEN,[||] LISA ALVAREZ-COHEN,[§,||] IVAN MORENO-ANDRADE,[⊥] CHUN-LIN HUANG,[§] AND SLAWOMIR W. HERMANOWICZ[§]

State Key Laboratory of Pollution Control and Resource Reuse, Tongji University, Shanghai 200092, PR China, Chinese Research Academy of Environmental Sciences, Beijing 100012, PR China, Department of Civil and Environmental Engineering, University of California, Berkeley, California 94720, Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, and Laboratory for Research on Advanced Processes for Water Treatment, Instituto de Ingeniería, Unidad Academica Juriquilla, Universidad Nacional Autonoma de Mexico, Blvd. Juriquilla 3001, Queretaro 76230, Mexico

Received May 7, 2010. Revised manuscript received August 1, 2010. Accepted August 17, 2010.

Current knowledge of the microbial communities within biological wastewater treatment reactors is incomplete due to limitations of traditional culture-based techniques and despite the emergence of recently applied molecular techniques. Here we demonstrate the application of high-density microarrays targeting universal 16S rRNA genes to evaluate microbial community composition in five biological wastewater treatment reactors in China and the United States. Results suggest a surprisingly consistent composition of microbial community structure among all five reactors. All investigated communities contained a core of bacterial phyla (53−82% of 2119 taxa identified) with almost identical compositions (as determined by colinearity analysis). These core species were distributed widely in terms of abundance but their proportions were virtually the same in all samples. Proteobacteria was the largest phylum and Firmicutes, Actinobacteria, Bacteroidetes were the subdominant phyla. The diversity among the samples can be attributed solely to a group of operational taxonomic units (OTUs) that were detected only in specific samples. Typically, these organisms ranked somewhat lower in terms of abundance but a few were present is much higher proportions.

Introduction

Bacteria are crucial to ecosystem functioning, and play vital roles in carbon, nitrogen and sulfur cycles (*1*). Biological wastewater treatment, utilization processes such as activated sludge (AS) or membrane biological reactors (MBRs), is used to protect water environments by removing contaminants from wastewater including organics, nitrogen, and phosphorus (*2, 3*). These processes are applied at very large scales (~10^5 m^3/d), possibly representing the largest applications of bioprocess engineering. Current knowledge of the microbial communities within biological wastewater treatment reactors is incomplete due to limitations of traditional culture-based techniques and despite the recent emergence of molecular techniques (*4, 5*). Microbial communities have been characterized in natural and engineered ecosystems using a variety of molecular methods. For example, ribosomal spacer analysis (RISA) (*6*), terminal restriction fragment length polymorphism (t-RFLP) (*7*), denaturing gradient gel electrophoresis (DGGE) (*8*), 16S rRNA clone libraries (*9*), and fluorescence in situ hybridization (FISH) (*10*) were applied to evaluate bacterial community structure not only in biological reactors from wastewater treatment plants but also in other ecosystems such as edaphic (*11*), marine (*12*) and atmospheric (*13*) microbial systems. However, the results obtained with these methods can be incomplete as they do not capture the whole complexity of microbial communities (*14*).

Microarray-based genomics is an emerging technology to study microbial communities (*15*). Microarrays are a powerful tool for viewing the expression of tens of thousands of genes simultaneously in a single experiment (*16*), rendering them a specific, sensitive, quantitative, and high-throughput tool for microbial detection, identification, and characterization in natural environments. This is the main technical advantage over other taxonomic nucleic acid-based assays, which are limited by the rate at which sequences can be analyzed, especially in complex samples. Specifically, high-density microarrays that target 16S rRNA genes can be used to taxonomically identify large numbers of organisms with no subsequent DNA isolation and sequencing. Their applications have been dramatically extended to environmental systems in recent years (*17*−*19*). However, we are not aware of any report using a high-density universal 16S rRNA miroarray to compare microbial communities between various wastewater treatment bioreactors operated at different geographic locations.

We propose that aerobic biological treatment processes have stable and similar bacterial community structures primarily influenced by the commonality of the influent wastewater. To evaluate this hypothesis, we applied high-density microarrays containing 506 944 probes targeting 8935 clusters of 16S rRNA genes (PhyloChips) to analyze the composition of bacterial communities from five different wastewater treatment bioreactors in China and the United States.

Materials and Methods

PhyloChip. As shown in Table 1, all five bioreactors treat domestic sewage but they vary significantly in their size, process configuration and operational parameters. Bench scale communities were sampled several months after their start and had stable microbial communities prior to the study. The pilot plants were operated for at least one year prior to analysis. Full-scale plants have been operational for many years.

Biomass samples were collected from five systems during stable operation. DNA was extracted in triplicate with a Fast DNA Spin Kit (Qiagen, CA) as described in the manufacturer's instructions. 16S rRNA gene fragments were amplified from

* Corresponding author phone: 86-21-65980440; fax: 86-21-65986313; e-mail: siqingxia@gmail.com.
† Tongji University.
‡ Chinese Research Academy of Environmental Sciences.
|| Lawrence Berkeley National Laboratory.
§ University of California, Berkeley.
⊥ Universidad Nacional Autonoma de Mexico.

10.1021/es101554m © 2010 American Chemical Society
Published on Web 09/02/2010

TABLE 1. Characteristics of Five Bioreactors in China and the U.S.[a]

name	location	types of bioreactor	scale	flow (m^3/day)	temp(°C)	influent water characteristics(mg/L)			
						types	NH_4^+	COD	TSS
China-1	Shanghai, China	CAS	plant	$(5.7-6.0) \times 10^5$	10	domestic	30−35	312−447	166−357
China-2	Shanghai, China	A/O-MBR	pilot	20−40	10	domestic	11−66	50−431	65−300
China-3	Shanghai, China	MBR	bench	0.024	10	synthetic	10−40	100−350	40−180
USA-1	San Francisco, U.S.	CAS	plant	$(0.3-2.2) \times 10^6$	8	domestic	14−33	78−695	57−242
USA-2	Berkeley, U.S.	MBR	bench	0.024	20	synthetic	30−40	152−210	0

[a] Note: CAS: conventional activated sludge; MBR: membrane bioreactor; A/O-MBR: anaerobic/oxic-membrane bioreactor; COD: chemical oxygen demand; TSS: total suspended solids.

the total DNA extracted from the samples with Taq DNA hot start polymerase (Promega Inc., WI) using universal bacterial primer sets 27f (AGRGTTTGATCMTGGCTCAG) and 1492R (GGTTACCTTGTTACGACTT). At least three replicate PCR reactions were performed per sample and pooled before analysis. The PCR mixture contained 1.25 U of Taq polymerase, 1× PCR buffer (Promega, WI), 2 mM $MgCl_2$, 0.5 μmol of each primer, 200 μM deoxynucleoside triphosphate, and 40 ng of template DNA. PCR amplifications were carried out in a total volume of 50 μL in 0.2 mL tubes using a DNA gradient thermocycler (Mastercycles, Eppendorf) to maximize bacterial diversity. A hot-start PCR program was used for all amplifications to minimize nonspecific amplification. The PCR were performed at a range of 12 different annealing temperatures between 48 and 58 °C. The following conditions were used for PCR: 2 min of initial denaturation at 95 °C and 25 cycles of 2 min at 95 °C, 45 s between 48 and 58 °C, 2 min at 72 °C, and the final extension for 10 min at 72 °C.

The amplicons from the 12 different annealing temperatures were combined. The sizes of PCR products were assessed by Agilent 2100 Bioanalyzer using DNA 7500 chips (Agilent, CA). The amplicon concentrations were measured with a NanoDrop 3300 Fluorospectrometer (Thermo Sci, DE) using PicoGreen assay for dsDNA (Invitrogen, CA). Then the bacterial PCR products were concentrated using Microcon YM100 spin filter (Microcon, MA). A total of 500 ng of PCR product was analyzed by PhyloChip.

G2 PhyloChips were used with probes designed in the Lawrence Berkeley National Laboratory (*20*). Each PhyloChip included 506 944 probes with 297 851 probes targeted 8935 clusters of 16S rRNA gene sequences. The remaining probes were used for image orientation, normalization controls, or for pathogen-specific signature amplicon detection using additional targeted regions of the chromosome (*17*). Each of the 8935 clusters, containing approximately 3% sequence divergence, was considered an operational taxonomic unit (OTU) representing all 121 demarcated prokaryotic orders. The taxonomic family of each OTU was assigned according to the placement of its member organisms in Bergey's Taxonomic Outline (*21*). The taxonomic outline was consulted for phylogenetic classes containing uncultured environmental organisms or unclassified families belonging to named higher taxa. The OTUs comprising each family were clustered into 842 subfamilies by transitive sequence identity according to a previously described method (*20*).

The chosen oligonucleotides were synthesized by a photolithographic method at Affymetrix Inc. (Santa Clara, CA) directly onto a 1.28 × 1.28 cm glass surface at an approximate density of 10 000 molecules/μm^2 to create the PhyloChips. The probe features were arranged as a grid of 712 rows and columns. Thus, each unique probe sequence on the array occupied a square with an 18 μm side and had a copy number of roughly 3.2×10^6.

The amplicon pools were fragmented to 50−200 bp using DNase I (0.02 U/mg DNA; Invitrogen) and then terminally labeled with biotin. Next, the labeled DNA was denatured (99 °C for 5 min) and hybridized to the DNA microarrays at 48 °C overnight at 60 rpm. PhyloChip washing and staining were performed according to standard Affymetrix protocols as described previously (*17, 21*).

Each PhyloChip was scanned and recorded as a pixel image, and individual signal values and intensities were completed using standard Affymetrix software (GeneChip Microarray Analysis Suite, version 5.1). Every OTU was interrogated by 24 replicate probes for precision control. The positive fraction (pf value) was calculated as the number of positive probe pairs (Supporting Information (SI) Materials and Methods) divided by the total number of probe pairs in each probe set (i.e., OTU). An OTU was considered "present" when its pf value was greater than 0.9.

Quantitative Analysis of Microbial Communities. The OTU abundance (measured by fluorescence intensity) was normalized by respective sample fluorescence maxima to yield the relative OTU abundance x ($0 < x < 1$). The OTUs were then divided into classes based on their relative abundances x yielding the frequency distributions $P(x)$. Similarity between the populations was quantified using the colinearity analysis. For this purpose, the composition of each sample is represented by an N-dimensional vector in the OTU space. The coordinates of such a vector are (OTU_1, OTU_2, ... OTU_N) where OTU_i is the relative abundance of i-th OTU in the sample and N is the number of OTUs detected in each sample. With such a representation of each microbial community, the degree of similarity can be expressed as the angle between two vectors, each corresponding to a particular community. A visual representation of this approach is shown in SI Figure S1 for three vectors (representing three populations) with only three species present for simplicity. In this hypothetical picture, samples A and B have very close compositions indicated by a small angle α while samples A and C have much diverse compositions since angle β is much larger. To assess the significance of our results we performed 500 Monte Carlo simulations in which the angles between random vectors were calculated. Each vector was composed of 1000 elements ($N = 1000$). The average value of the angle between the random vectors was 41.43° (0.723 rad) with the standard deviation of 0.89° (0.016 rad).

Results and Discussion

The Composition of Bacterial Community. Using the PhyloChip, 2119 distinct operational taxonomic units (OTUs) were detected at least in one of the five samples. Individual samples contained somewhat smaller number of OTUs: 1729, 1546, 1500, 1440, and 1126 in China-1, China-2, China-3, USA-1, and USA-2, respectively. Thus, the individual samples contained between 53 and 82% of the total OTU ensemble. While the number of OTUs in the individual samples varied, all five samples contained a core of 859 OTUs, indicating a large degree of similarity among the five samples.

Based simply on the presence or absence of specific OTUs, Proteobacteria was the predominant phylum, constituting between 50 and 62% of all detected OTUs (Figure 1A). Firmicutes, Actinobacteria, and Bacteroidetes were the subdominant groups, each comprising between 5 and 18%

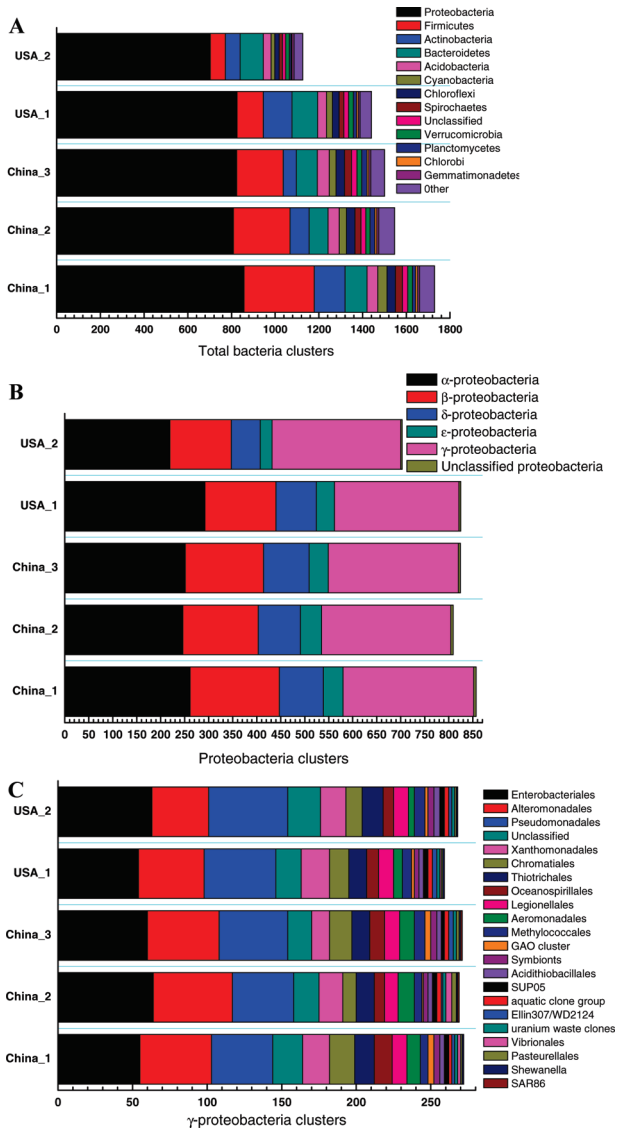

FIGURE 1. Bacterial community composition in five samples. a, the community composition of total bacteria grouped by phyla. Some phyla were assigned to "other" if detected in fewer than eight OTUs. b, the community composition of the phylum Proteobacteria. c, the community composition of the class γ-Proteobacteria.

of the detections. These four bacterial groups represented approximately 80% of bacteria detected within the five samples. This similarity, based on the presence/absence of OTUs, extended down to more specific taxa. For example, within Proteobacteria the γ subdivision was the largest group (31−38%) closely followed by α-Proteobacteria (from 30 to 35%) (Figure 1B). Within the γ-Proteobacteria, 22 taxa were identified (Figure 1C) with Enterobacteriales being the dominant group within a narrow range of 20−25% of all five samples. They were followed in dominance by Alteromonadales and Pseudomonadales, each also constituting a similar fraction in each population (14−19% and 15−20%, respectively). The seven other detected groups (aquatic clone group, Ellin307/WD2124, uranium waste clones, Vibrionales, Pasteurellales, Shewanella, SAR86) had many fewer detections (1−3 OTUs) in all samples and constituted about 7% of γ-Proteobacteria. We refer here to bacterial groups rather than species. Each group potentially contains more than one species (*20, 21*).

Even at the more distinct level of taxa, the samples from different bioreactors had similar composition. For example, Acidobacteria were represented by 36−51 OTUs (constituting 2.8−3.6% of the detected OTUs), while only eight or nine OTUs of Gemmatimonadetes were found in the samples (0.5−0.8% of total OTUs).

The remaining Proteobacteria subdivisions (α, β, δ, ε, and unclassified) had fewer detections but were also present in all samples in very similar proportions (Figure 2). It is interesting that β-Proteobacteria constituted only between 18 and 20% of total Proteobacteria detections despite being identified commonly as a main subpopulation in wastewater treatment bioreactors (*22*) (Figure 1B). Perhaps this discrepancy can be explained by the fact that PhyloChips are able to detect a much larger variety of microorganisms leading to a more complete characterization of complex populations (*21*).

In more general terms, while molecular tools such as RISA, t-RFLP, DGGE, 16S rRNA clone libraries and FISH are extremely useful for characterizing microbial community composition, they are generally not effective for detecting the depth of the structure of highly complex communities. For example, although PCR-DGGE can generally detect bacterial groups larger than 1% of the population, single bands sometimes do not correspond to single bacterial species (*23*). The T-RFLP results are typically limited to only fifty or so most abundant organisms (*24*), so it cannot effectively determine phylogenetic richness in extremely complex communities (*25*). Finally, 16S rRNA clone libraries are also subject to limitations of low clone numbers resulting in low-sensitivity, with even libraries of greater than 1000 clones demonstrating only moderate sensitivity in complex communities that miss many rare taxa (*5, 21*). In contrast, 454-pyrosequencing can be used to sequence the metagenome of complex communities and is more specific and has significantly higher throughput than the PhyloChip. However, it has the disadvantage of being extremely more expensive than most of the alternative technologies. The current cost and complications associated with pyrosequencing complex metagenomes at appropriate read levels is likely to limit its use for understanding the complex ecology of wastewater treatment bioreactors (*26, 27*). Previous researchers acquired 378 601 sequences with an average read length of 250.4 bp from the activated sludge basin of a wastewater treatment plant in Charlotte using 454-pyrosequencing, but were able to assemble only 0.3% of the sequences into significant contigs (*27*), significantly hindering data interpretation.

Despite the distinct advantages of the PhyloChip for providing rapid and inexpensive microbial community profiling, it also has disadvantages. For example, only targeted (already known) OTUs tiled onto the microarray are detected. In addition, with amplification of DNA using bacterial primers, there is a potential for PCR bias (despite the 12 different annealing temperatures) and only the bacterial community is queried whereas the archaeal and fungal populations are ignored. Consequently, in this study, we do not know if the biological treatment systems from the five different locations harbored similar archaeal and fungal populations.

Quantitative Analysis. To quantify the abundance of OTUs (as opposed to merely presence/absence) in each sample, the average hybridization (fluorescence) intensity was measured. To control for intensity variations from sample to sample, the measured intensities were normalized by the maximum intensity value in each sample and assumed to be equal to the fraction of the corresponding DNA extracted from the population. The frequency distributions $P(x)$ approximately followed beta probability distribution as expected for normalized variables (*28*) (SI Figure S2). The only significant deviations from the theoretical distributions ("fat tails") were for a few most abundant OTUs (top 1−3%

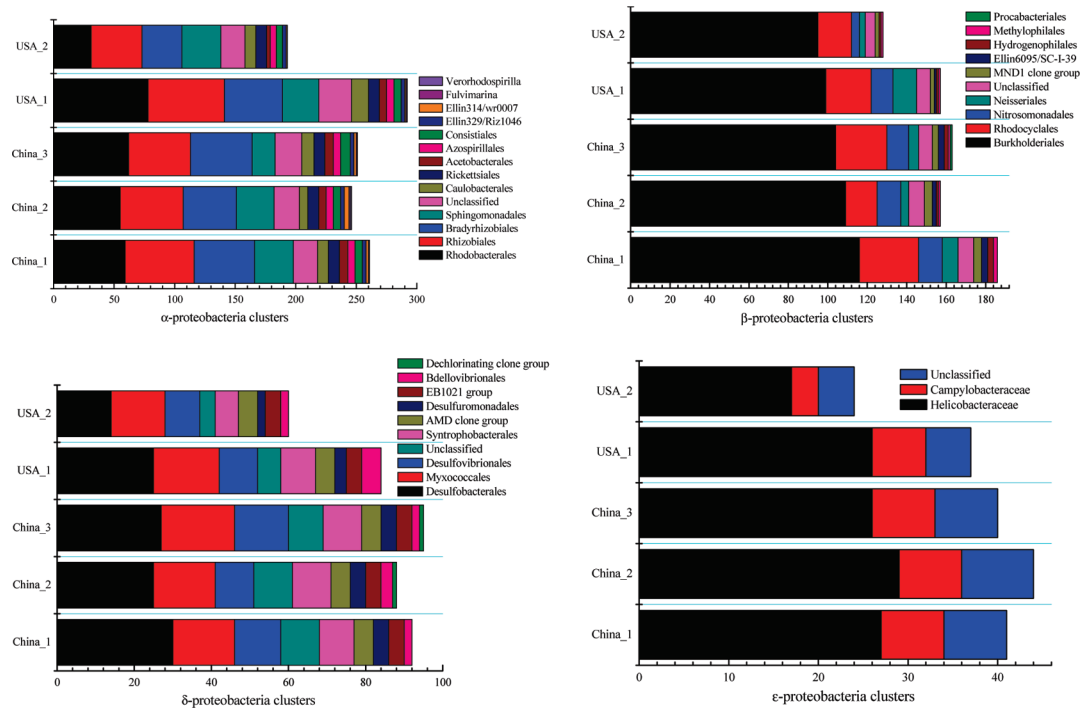

FIGURE 2. Bacterial community composition of α-, β-, δ-, and ε-Proteobacteria.

TABLE 2. Pair-Wise Common OTU Numbers

population (sample)	shannon index H'	number of OTUs common for both populations (samples) China-1	China-2	China-3	USA-1	USA-2
China-1	7.4		1378	1365	1247	995
China-2	7.3			1320	1171	926
China-3	7.3				1165	961
USA-1	7.2					1014
USA-2	7.0					

of the population, 10−50 OTUs) which were over-represented compared to the beta distribution.

To assess the internal (within-sample) complexity of individual microbial populations, Shannon diversity index H' values were calculated (*29*). The values of H' were quite close across the five samples ranging from 7.0 for USA-2 to 7.4 for China-1 (Table 2). These H' values are typical for diverse microbial populations without a few strongly dominant taxa.

Further analysis was performed to quantify the similarity between microbial populations in the five samples. For this purpose, a colinearity analysis was used. For the purpose of the analysis, the vector representing each of the five populations (samples) has to be embedded in the *N*-dimensional space. We made two choices of *N*. The first choice was to use the common core of $N = 859$ OTUs present in all five samples. Based on this choice, the angles between each pair of five vectors were calculated. The angle between the vectors then represents a measure of "relatedness/similarity" such that perfectly identical populations would be described by a zero angle while completely different populations would be represented by orthogonal (90°) vectors. Consequently, populations with a 1° angle would be considered almost identical, whereas an 85° angle would correspond to highly dissimilar populations. The calculated values among the five population samples were quite small, between 4.7° and 13.8° (certainly significantly different from random vectors) indicating a very strong similarity of microbial populations in all five samples (Table 3). To further assess the effects of

TABLE 3. Microbial Composition Similarity for Common Core of OTUs (N = 859 OTUs)[a]

population (sample)	angle between two populations (degrees) China-1	China-2	China-3	USA-1	USA-2
China-1	0	8.2	4.7	12.7	13.5
China-2		0	7.5	13.1	12.5
China-3			0	12.3	13.8
USA-1				0	11.4
USA-2					0

[a] Note: The diagonal values are zero by definition.

TABLE 4. Microbial Composition Similarity for Most Abundant Core OTUs

number of top core OTUs	minimum angle (degrees)	maximum angle (degrees)
10	2.9	6.8
20	2.7	10.2
50	2.6	8.8
100	2.8	9.3
200	3.1	9.5
400	3.3	10.3
859 (whole core)	4.7	13.8

dominating OTUs, a similar analysis was performed for each pairwise combination of populations but only with a set of most abundant OTUs selected from the common core. We chose between 10 and 400 top OTUs and calculated the angles for these subsets. The results (Table 4) show that the similarity (characterized by very small angles and almost colinear vectors) of the composition was not dominated only by the most abundant members but extended also to less abundant members in the core. As an example, SI Table S1 lists the top 20 OTUs in the China-1 sample and the corresponding ranks and relative abundance of these OTUs in other samples. Although the ranks do not match perfectly, the members that are abundant in one sample are also abundant in other samples.

The second choice of N was the total ensemble of $N = 2119$ OTUs detected in at least in one sample. If a particular OTU was not detected in a sample its abundance was assumed to be zero with the corresponding vector coordinates also set to zero. This approach resulted in much larger pairwise angles between the vectors, obviously because of larger intersample diversity. The angles varied from 25° (between China-2 and China-3) to 40° (between China-1 and USA-2) (SI Table S2). However, even these larger angles were significantly lower (and statistically different) from random vectors.

To analyze further the effects of undetected OTUs on the diversity among the populations, we compared each pair from the five populations. For each pairwise comparison, we used the largest set of OTUs common to each pair. The N values were obviously larger than the $N = 859$ for the core, smaller than the total number $N = 2119$ of detected OTUs and varied between 926 and 1378 (Table 2). The resulting angles in this pairwise comparison were very small and virtually the same as those based on the common core of 859 OTUs, despite inclusion of up to 60% more OTUs (SI Table S3).

Thus, the diversity between microbial populations sampled in the five bioreactors is almost completely due to taxa that were present in some samples but were not detected in others. When the populations are compared based upon any subset of taxa present in more than one sample, their compositions were virtually identical. It appears that the microbial populations in these bioreactor samples consist of a common core of taxa (one-half to two-thirds of OTUs) with a very similar composition, with the remaining portion of diverse taxa detected solely in one specific population. SI Table S4 lists the 10 most abundant taxa unique to each sample.

We further analyzed the core and the remainder of the population for relative abundance of microbial taxa in each. We found that in each sample, no core taxa were present in the least abundant quintile. However, the OTUs unique to each sample were also found at higher levels, up to 80% by abundance. From this point of view, each population was composed (approximately) of the top quintile that contained only core taxa (common for all samples), the bottom quintile that contained only unique taxa, and the remaining middle three quintiles (20−80%) that contained core and unique taxa.

The diversity among the samples can be attributed solely to a group of unique OTUs that were detected only in specific samples. Typically, these taxa ranked somewhat lower in terms of abundance (bottom half), but a few were present is much higher proportions. It is not clear at this time which of the factors contributes to the formation of the common population core and which are responsible for the unique set of taxa in each sample. It seems that there was stronger commonality between the samples from China than between the Chinese and U.S. samples, although the difference is not very large. Siripong and Rittmann also showed similar community structure in all seven full-scale municipal wastewater treatment plants but only for nitrifying bacteria (*30*). It is possible that the characteristics of wastewater and the origin of inoculum may lead to commonalities in the population.

Acknowledgments

We acknowledge the support provided by the national high technology research and development program (863), the national natural science foundation, and the national key project of water pollution control of China (2008ZX07208-001&003). L.D. was supported by China Scholarship Council. Ivan Moreno Andrade was supported by a postdoctoral fellowship from the UC Mexus program. Chun-Lin Huang was supported by a grant from Taiwan.

Supporting Information Available

Figures S1−S3 and Tables S1−S4. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) Strom, S. L. Microbial ecology of ocean biogeochemistry: A community perspective. *Science* **2008**, *320*, 1043–1045.
(2) Horner-Devine, M. C.; Martiny, A. C. Biogeochemistry—News about nitrogen. *Science* **2008**, *320*, 757–758.
(3) Ahmed, Z.; Lim, B. R.; Cho, J.; Song, K. G.; Kim, K. P.; Ahn, K. H. Biological nitrogen and phosphorus removal and changes in microbial community structure in a membrane bioreactor: Effect of different carbon sources. *Water Res.* **2008**, *42*, 198–210.
(4) Green, B. D.; Keller, M. Capturing the uncultivated majority. *Curr. Opin. Biotechnol.* **2006**, *17*, 236–240.
(5) Fuhrman, J. A. Microbial community structure and its functional implications. *Nature* **2009**, *459*, 193–199.
(6) Smith, N. R.; Yu, Z. T.; Mohn, W. W. Stability of the bacterial community in a pulp mill effluent treatment system during normal operation and a system shutdown. *Water Res.* **2003**, *37*, 4873–4884.
(7) Eschenhagen, M.; Schuppler, M.; Roske, I. Molecular characterization of the microbial community structure in two activated sludge systems for the advanced treatment of domestic effluents. *Water Res.* **2003**, *37*, 3224–3232.
(8) Whiteley, A. S.; Bailey, M. J. Bacterial community structure and physiological state within an industrial phenol bioremediation system. *Appl. Environ. Microbiol.* **2000**, *66*, 2400–2407.
(9) Duan, L.; Moreno-Andrade, I.; Huang, C. L.; Xia, S. Q.; Hermanowicz, S. W. Effects of short solids retention time on microbial community in a membrane bioreactor. *Bioresour. Technol.* **2009**, *100*, 3489–3496.
(10) Luxmy, B. S.; Nakajima, F.; Yamamoto, K. Analysis of bacterial community in membrane-separation bioreactors by fluorescent in situ hybridization (FISH) and denaturing gradient gel electrophoresis (DGGE) techniques. *Water Sci. Technol.* **2000**, *41*, 259–268.
(11) Monson, R. K.; Lipson, D. L.; Burns, S. P.; Turnipseed, A. A.; Delany, A. C.; Williams, M. W.; Schmidt, S. K. Winter forest soil respiration controlled by climate and microbial community composition. *Nature* **2006**, *439*, 711–714.
(12) Webster, G.; Newberry, C. J.; Fry, J. C.; Weightman, A. J. Assessment of bacterial community structure in the deep sub-seafloor biosphere by 16S rDNA-based techniques: a cautionary tale. *J. Microbiol. Methods* **2003**, *55*, 155–164.
(13) Maron, P. A.; Mougel, C.; Lejon, D. P. H.; Carvalho, E.; Bizet, K.; Marck, G.; Cubito, N.; Lemanceau, P.; Ranjard, L. Temporal variability of airborne bacterial community structure in an urban area. *Atmos. Environ.* **2006**, *40*, 8074–8080.
(14) Sanz, J. L.; Kochling, T. Molecular biology techniques used in wastewater treatment: An overview. *Process Biochem.* **2007**, *42*, 119–133.
(15) Zhou, J. H. Microarrays for bacterial detection and microbial community analysis. *Curr. Opin. Biotechnol.* **2003**, *6*, 288–294.
(16) Wilson, K. H.; Wilson, W. J.; Radosevich, J. L.; DeSantis, T. Z.; Viswanathan, V. S.; Kuczmarski, T. A.; Andersen, G. L. High-density microarray of small-subunit ribosomal DNA probes. *Appl. Environ. Microbiol.* **2002**, *68*, 2535–2541.
(17) Brodie, E. L.; DeSantis, T. Z.; Joyner, D. C.; Baek, S. M.; Larsen, J. T.; Andersen, G. L.; Hazen, T. C.; Richardson, P. M.; Herman, D. J.; Tokunaga, T. K.; et al. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl. Environ. Microbiol.* **2006**, *72*, 6288–6298.
(18) Huyghe, A.; Francois, P.; Charbonnier, Y.; Tangomo-Bento, M.; Bonetti, E. J.; Paster, B. J.; Bolivar, I.; Baratti-Mayer, D.; Pittet, D.; Schrenzel, J.; et al. Novel Microarray design strategy to study complex bacterial communities. *Appl. Environ. Microbiol.* **2008**, *74*, 1876–1885.
(19) Gentry, T. J.; Wickham, G. S.; Schadt, C. W.; He, Z.; Zhou, J. Microarray applications in microbial ecology research. *Microbial Ecol.* **2006**, *52*, 159–175.
(20) DeSantis, T. Z.; Dubosarskiy, I.; Murray, S. R.; Andersen, G. L. Comprehensive aligned sequence construction for automated design of effective probes (CASCADE-P) using 16S rDNA. *Bioinformatics* **2003**, *19*, 1461–1468.
(21) DeSantis, T. Z.; Brodie, E. L.; Moberg, J. P.; Zubieta, I. X.; Piceno, Y. M.; Andersen, G. L. High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microbial Ecol.* **2007**, *53*, 371–383.

(22) Snaidr, J.; Amann, R.; Huber, I.; Ludwig, W.; Schleifer, K. H. Phylogenetic analysis and in situ identification of bacteria in activated sludge. *Appl. Environ. Microbiol.* **1997**, *63*, 2884–2896.
(23) Muyzer, G.; Smalla, K. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie van Leeuwenhoek* **1998**, *73*, 127–141.
(24) Sakano, Y.; Kerkhof, L. Assessment of changes in microbial community structure during operation of an ammonia biofilter with molecular tools. *Appl. Environ. Microbiol.* **1998**, *64*, 4877–4882.
(25) Dunbar, J.; Ticknor, L. O.; Kuske, C. R. Assessment of microbial diversity in four southwestern United States soils by 16S rRNA gene terminal restriction fragment analysis. *Appl. Environ. Microbiol.* **2000**, *66*, 2943–2950.
(26) Hudson, M E. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Resour.* **2008**, *8*, 3–17.
(27) Sanapareddy, N.; Hamp, T. J.; Gonzalez, L. C.; Hilger, H. A. Molecular Diversity of a North Carolina Wastewater Treatment Plant as Revealed by Pyrosequencing. *Appl. Environ. Microbiol.* **200975**, 1688–1696.
(28) Nadarajah, S.; Gupta, A. K. Characterizations of the beta distribution. *Commun. Stat-Theory Methods* **2004**, *33*, 2941–2957.
(29) Shanno, C. R. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, 1963.
(30) Siripong, S; Rittmann, B. E. Diversity study of nitrifying bacteria in full-scale municipal wastewater treatment plants. *Water Res.* **2007**, *41*, 1110–1120.

ES101554M