

# Lawrence Berkeley National Laboratory

## LBL Publications

### Title

Effects of operational parameters on bacterial communities in Hong Kong and global wastewater treatment plants.

### Permalink

<https://escholarship.org/uc/item/1pz7x888>

### Journal

mSystems, 9(3)

### Authors

Zhang, Yulin

Deng, Yu

Wang, Chunxiao

et al.

### Publication Date

2024-03-19

### DOI

10.1128/msystems.01333-23

Peer reviewed

# Effects of operational parameters on bacterial communities in Hong Kong and global wastewater treatment plants

Yulin Zhang,<sup>1</sup> Yu Deng,<sup>1</sup> Chunxiao Wang,<sup>1</sup> Shuxian Li,<sup>1</sup> Frankie T. K. Lau,<sup>2</sup> Jizhong Zhou,<sup>3</sup> Tong Zhang<sup>1,4</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 16.

**ABSTRACT** Wastewater treatment plants (WWTPs) are indispensable biotechnology facilities for modern cities and play an essential role in modern urban infrastructure by employing microorganisms to remove pollutants in wastewater, thus protecting public health and the environment. This study conducted a 13-month bacterial community survey of six full-scale WWTPs in Hong Kong with samples of influent, activated sludge (AS), and effluent to explore their synchronism and asynchronism of bacterial community. Besides, we compared AS results of six Hong Kong WWTPs with data from 1,186 AS amplicon data in 269 global WWTPs and a 9-year metagenomic sequencing survey of a Hong Kong WWTP. Our results showed the compositions of bacterial communities varied and the bacterial community structure of AS had obvious differences across Hong Kong WWTPs. The co-occurrence analysis identified 40 pairs of relationships that existed among Hong Kong WWTPs to show solid associations between two species and stochastic processes took large proportions for the bacterial community assembly of six WWTPs. The abundance and distribution of the functional bacteria in worldwide and Hong Kong WWTPs were examined and compared, and we found that ammonia-oxidizing bacteria had more diversity than nitrite-oxidizing bacteria. Besides, Hong Kong WWTPs could make great contributions to the genome mining of microbial dark matter in the global “wanted list.” Operational parameters had important effects on OTUs’ abundance, such as the temperature to the genera of *Tetrasphaera*, *Gordonia* and *Nitrospira*. All these results obtained from this study can deepen our understanding of the microbial ecology in WWTPs and provide foundations for further studies.

**IMPORTANCE** Wastewater treatment plants (WWTPs) are an indispensable component of modern cities, as they can remove pollutants in wastewater to prevent anthropogenic activities. Activated sludge (AS) is a fundamental wastewater treatment process and it harbors a highly complex microbial community that forms the main components and contains functional groups. Unveiling “who is there” is a long-term goal of the research on AS microbiology. High-throughput sequencing provides insights into the inventory diversity of microbial communities to an unprecedented level of detail. At present, the analysis of communities in WWTPs usually comes from a specific WWTP and lacks comparisons and verification among different WWTPs. The wide-scale and long-term sampling project and research in this study could help us evaluate the AS community more accurately to find the similarities and different results for different WWTPs in Hong Kong and other regions of the world.

**KEYWORDS** wastewater treatment plants, activated sludge, nitrification, assembly mechanism, operational parameters, microbial dark matter

Wastewater treatment plants (WWTPs) are an indispensable component of modern cities, as they can remove pollutants in wastewater to prevent anthropogenic activities (1). Activated sludge (AS) (2) is a fundamental wastewater treatment process

**Editor** Thulani P. Makhhalanyane, University of Pretoria, Pretoria, South Africa

Address correspondence to Tong Zhang, zhangt@hku.hk.

Yulin Zhang and Yu Deng contributed equally to this article. Author order was determined in order of contribution to this research article.

The authors declare no conflict of interest.

See the funding table on p. 16.

**Received** 13 December 2023

**Accepted** 26 January 2024

**Published** 27 February 2024

Copyright © 2024 Zhang et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

and it harbors a highly complex microbial community that forms the main components and contains functional groups. Unveiling “who is there” is a long-term goal of the research on AS microbiology. Currently, high-throughput sequencing (HTS) provides insights into the inventory diversity of microbial communities to an unprecedented level of detail and enables the detection of more taxa including a large percentage of uncultured organisms.

Studies of compositions of microbial communities have increased researchers’ understanding of the associations and relationships between microorganisms in both natural and engineering systems (3, 4). For example, seasonal variation had a big impact on the AS bacterial community (1, 2) and the biodegradability of the influent could determine WWTP performance by affecting the microbial community of AS (5). Industrial AS exhibits a unique profile that is distinct from the characteristics of municipal AS due to the composition difference between industrial and municipal wastewater (6). Apart from the overall microbial communities, the co-occurrence patterns of microbial communities will provide new insights into the AS system (1). Network analysis is a useful tool for studying microbial associations and has been applied to determine microbial coexistence in complex communities, such as soils (7) and oceans (8). The application of such network analysis on the AS community may improve our understanding of potential microbial relationships for bacterial adjustment (9) and provide ideas to improve the treatment efficiency of AS (2, 10, 11). Understanding the assembly rules like neutral and niche theories (12) and finding corresponding assembly driving factors of the microbial community are also important ecological problems. Both stochastic and deterministic processes operate concurrently for WWTPs (13). For instance, researchers revealed that microbial communities in AS were dominated by deterministic processes with temperature as the most important driver (14, 15). However, a global AS survey (16) and an AS study in China (17) showed more crucial positions of stochastic processes in community assembly (18).

One of the most important functions of WWTPs is to remove nutrients such as nitrogen and phosphorus from sewage (1). These important processes are performed by functional groups such as ammonia-oxidizing bacteria (AOB), nitrite-oxidizing bacteria (NOB) and polyphosphate-accumulating organisms (PAOs) (19), and can be affected by bulking and foaming bacteria (BFB). A comprehensive understanding of functional groups will provide valuable insight into the performance evaluation and routine maintenance of WWTPs. Genome sequencing enhances our understanding of the biological world by providing blueprints for the evolutionary and functional diversity that shapes the biosphere (20). However, microbial genomes that are currently available are of limited phylogenetic breadth due to our inability to culture many environmental microorganisms (21). These unknown microbes whose genomes remain uncharacterized are colloquially called “microbial dark matter” (MDM) (22). The proportions of MDM in WWTPs are of great value to define their roles and a study in 2023 conducted an accurate survey of the MDM conditions in global WWTPs to propose a “wanted list” (23). We should evaluate the MDM conditions of Hong Kong WWTPs and explore their potential contributions to the MDM exploration for global WWTPs.

AS is an engineered ecosystem whose microbial communities originate from inoculation sludge that is purposefully chosen and could be affected by various factors, such as geographical locations, seasonal (environmental) variations (24), influent wastewater characteristics (25) and operational parameters (26). The rapid increase in both the quantity and complexity of daily operational data in the field of environmental engineering demands advanced data analysis approaches like redundancy analysis (RDA) to find or predict key parameters for bacteria (27). However, most earlier studies evaluated the effects of parameters on the whole AS microbiota and did not provide feedback for specific genera.

At present, the analysis of communities in WWTPs usually comes from a specific WWTP and lacks comparisons and verification among different WWTPs. The wide-scale and long-term sampling project could help us evaluate the AS community more

accurately to find synchronous and asynchronous results for different WWTPs in Hong Kong and other regions of the world. In this research, samples (influent, AS, and effluent) were monthly collected from six full-scale WWTPs in Hong Kong for 13 months to conduct the 16S rRNA gene amplicon sequencing, and Hong Kong AS data were compared with worldwide WWTP results to address the following ecological questions of the bacterial community in AS: (i) bacterial community conditions for six Hong Kong WWTPs, such as compositions, co-occurrence patterns and assembly mechanisms; (ii) abundance and distribution of functional groups in Hong Kong and global WWTPs; (iii) MDM conditions in Hong Kong and global WWTPs; and (iv) effects of operational parameters on OTUs' abundance in Hong Kong and global WWTPs. This investigation will advance our knowledge of the microbial ecology in AS systems to lay the groundwork for future investigations into the community regulation approach for WWTPs.

## MATERIALS AND METHODS

### WWTP, sampling, and sample pretreatment information

The influent (IN), AS, and effluent (EFF) samples were collected from six secondary WWTPs that applied biological treatment process in Hong Kong, including Shatin (ST), Tai Po (TP), Shek Wu Hui (SWH), Yuen Long (YL), Sai Kung (SK), and Stanley (STL) WWTP. All these WWTPs adopted the conventional AS process for carbon and nitrogen removal except for STL WWTP, which also applies the integrated fixed film activated sludge process (Table S1). Among the six WWTPs, ST WWTP is the largest WWTP in Hong Kong treating 260,000 m<sup>3</sup> of sewage from over 650,000 citizens per day. Besides, ST WWTP and TP WWTP treat sewage with a salinity of ~1% since Hong Kong uses seawater for toilet flushing with stable performance. Although the compositions of influent vary among WWTPs, the performances of the six WWTPs were generally stable throughout the year and the effluent of all WWTPs always meets the standard according to the record of Hong Kong DSD. Their technical and operational details are shown in Tables S1 and S2, respectively.

Individual samples were taken monthly from January 2018 to January 2019 in WWTPs and finally, we got 74 influent, 74 AS, and 74 effluent samples since the sampling work of SK WWTP was suspended following the damage caused by a super typhoon in September 2018. An amount of 300 mL of influent (mixing with 300 mL of 100% ethanol on site), 25 mL of AS (mixing with 25 mL of 100% ethanol on site), and 5 L of effluent (from the final sedimentation tank) was taken on-site in capped containers and transported to the laboratory immediately. Then, 600 mL of influent samples was centrifuged at 4750 rpm for 10 min (Beckman Coulter Avant J-15R, USA) to remove the supernatant and keep concentrated microorganisms, and then the remaining samples were further centrifuged at 15,000 × *g* for 3 min (Beckman Coulter Microfuge 20R, USA) to collect pellets for DNA extraction. An amount of 2 mL of AS samples was centrifuged at 15,000 × *g* for 3 min (Beckman Coulter Microfuge 20R, USA) to collect pellets for DNA extraction. 5 L of effluent samples was filtered with a 0.45-μm cellulose acetate membrane (Advantec MFS, Inc., USA) to collect the biomass for DNA extraction.

### DNA extraction, amplicon sequencing, and data processing

The genomic DNA was extracted by the FastDNA SPIN Kit for Soil (MPBiomedicals, France) according to the protocol of the manufacturer. The quality and quantity of extracted DNA were determined by NanoDrop Spectrophotometer ND-100 (ThermoFisher Scientific, USA) and Qubit 2.0 fluorometer (Life Technologies, USA), respectively. 300 bp paired-end sequencing reads were generated on a HiSeq (Illumina, San Diego, CA, USA) using the V3-V4 PCR primer set of 341F ACTCCTACGGGAGGCAGCAG and 806R GGACTACHVGGGTWTCTAAT by Beijing Genomics Institute (BGI, China). The raw data of 16S rRNA reads have been submitted to the National Center for Biotechnology Information (NCBI) with the BioProject ID [PRJNA1012295](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1012295).

DADA2 (28) in QIIME2 (version 2020.2) (29) was used to correct DNA sequences by establishing a customized sequence error mode based on the input sequences after the sequence quality control. The singletons were removed to guarantee sequence quality, and then the sequence number was normalized to 25,000 (Fig. S1) followed by the OTU generation with vsearch (30) at a 97% similarity level. The Silva 138 database (31) was applied to annotate the taxonomy information and the functional groups such as AOB, NOB, PAO and BFB groups were extracted by taxonomy annotation according to Table S3.

### Basic information and data analysis of global WWTPs

The worldwide WWTP survey contained 1,186 AS samples from 269 WWTPs. The information details of worldwide WWTPs and generation steps of OTUs were described in the previous study (16). The representative sequences were aligned by MAFFT (version 7.505) (32) for constructing the phylogenetic tree with FastTree2 (version 2.1.10) (33) and visualizing by iTOL (version 6) (34). The functional groups of global WWTPs were extracted by taxonomy annotation with the Silva 138 database (31) according to Table S3. AOB and NOB OTUs of Hong Kong and global WWTPs were aligned with the Genome Taxonomy Database (GTDB R207) (35) by BLASTn (36) with the top-hit result to determine the genome information by a similarity cutoff of 97%.

### Neutral community model for communities

The relationship between OTU detection frequency and their relative abundance within the community was predicted using the neutral community model (NCM) in R (version 4.3.1), which was used to assess the potential significance of stochastic processes on community assembly. The model employed in this study followed the methodology developed by Chen et al. (37) and enabled scholars to assess the crucial ecological processes that were challenging to detect directly but had a significant impact on the microbial community, such as dispersal and drift. In the analysis,  $Nm$  determined the correlation between occurrence frequency and regional relative abundance and  $R^2$  represented the overall fit to the neutral model. Using 1,000 bootstrap replicates, the 95% confidence intervals around all fitting statistics were calculated.

### Co-occurrence network creation and analysis

The co-occurrence of species-species associations was analyzed with the calculation of all pair-wise Spearman's rank coefficients ( $P < 0.05$ ) among OTUs that had  $>0.01\%$  average relative abundance in samples and visualized with Gephi (version 0.9) (38). Meanwhile, networks were also constructed by the `spiec.easi()` function from the R package of SpiecEasi (version 1.1.1) (39) with the default settings of the "mb" parameter to estimate the conditional dependence of each pair of OTUs. Relationships of species identified by both the Spearman ( $P < 0.01$ ) and SpiecEasi were conserved for the following co-occurrence analysis among different WWTPs in Hong Kong. The core (showed in  $>80\%$  of samples with a relative abundance  $>0.1\%$ ) and functional OTUs were used to further filter species relationships. Finally, 40 pairs of relationships were obtained (Fig. 2; Table S4) and visualized with Gephi (version 0.9) (38).

### MDM conditions, operational parameters and long-term data of Hong Kong WWTPs

The proportions of the genome-sequenced cells and taxa were calculated according to the method described in the published research (23).  $P_{\text{number}}$  and  $P_{\text{abundance}}$  were used to represent the genome-sequenced proportions of taxa (total OTU number) and corresponding cell number (total relative abundance sum). In summary, sequences of Hong Kong WWTPs were aligned as query via BLASTn (E value  $<1e^{-5}$ ) (36) with genomes in GTDB R207 to pick the top-hit sequence meeting the standard of 100% alignment and 100% identity and calculate the  $P_{\text{number}}$  and  $P_{\text{abundance}}$ . To examine the distribution of the global "wanted list" in Hong Kong, sequences of Hong Kong WWTPs were aligned

as query via BLASTn (E value  $<1e^{-5}$ ) (36) with sequences in the global “wanted list” (23) to pick the top-hit sequence meeting the standard of 100% alignment and 100% identity. The operational parameters of WWTPs were provided by Hong Kong DSD, and RDA was conducted by R (version 4.3.1). The long-term metagenomic sequencing data of ST AS (over 9 years) was collected from the former research (40) to extract the wanted metagenome-assembled genomes (MAGs) according to their taxonomic annotations.

## RESULTS AND DISCUSSION

### The conditions of the bacterial community for Hong Kong WWTPs

#### *The composition, diversity and dynamics of the bacterial community for Hong Kong WWTPs*

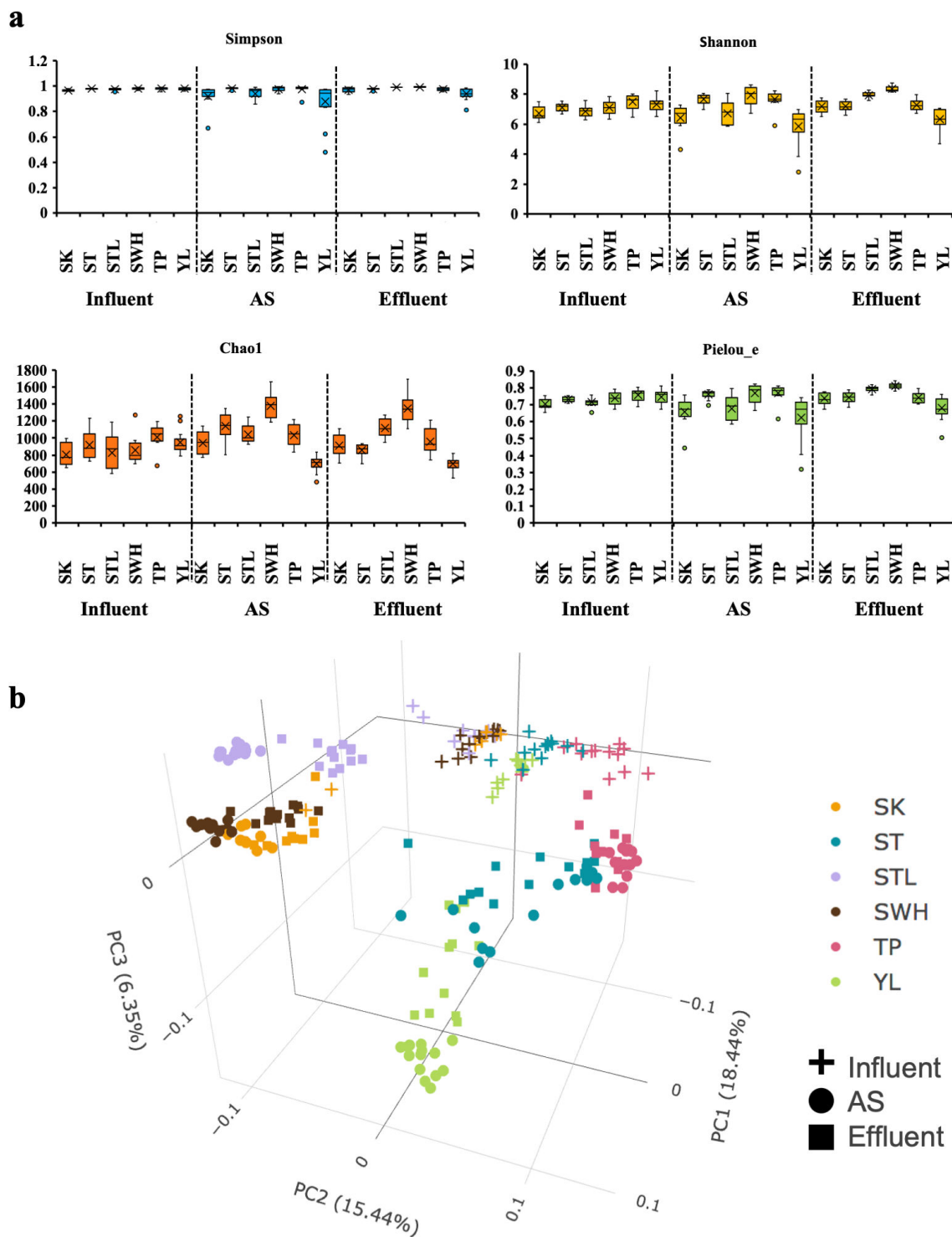
In total, AS samples of six Hong Kong WWTPs generated 7,649 OTUs with 48 phyla of bacteria (7,569 OTUs), 3 phyla of archaea (10 OTUs), and 70 OTUs with unclassified bacterial phyla (Fig. S2). Consistent with previous studies (1, 2), the phylum abundance of communities varied in different WWTPs but the dominant phyla in AS were the same, which were *Proteobacteria*, *Bacteroidetes*, *Chloroflexi*, *Patescibacteria*, *Actinobacteria*, *Acidobacteria*, and *Nitrospirae*. The performance of WWTPs could be more stable if AS communities have a higher diversity (5). Overall, the bacterial diversity of AS was relatively stable throughout the year, which was consistent with the stable performance of these six full-scale WWTPs (Fig. 1a; Fig. S3). The bacterial diversity of AS for different WWTPs varied greatly, and SWH and YL WWTPs had the highest and lowest richness and diversity, respectively. AS communities showed higher richness and evenness than influent and effluent communities in most WWTPs, while the diversity of effluent communities displayed notable fluctuations as they had the largest standard deviation among the three types of samples (Fig. S3).

The Bray-Curtis dissimilarity distance and principal coordinates analysis (PCoA) plot based on the relative abundance of bacteria at the phylum level was applied to calculate and distinguish the compositional differences and community relationships among WWTPs samples. Fig. 1b showed that the influent from different WWTPs had similar community compositions at the phylum level when compared with other types of samples (i.e., AS and effluent). This similarity can be attributed to the fact that the bacteria in influent mainly originated from the human gut (41, 42) and sewer pipes (43), and Hong Kong had a limited geographical area with similar dietary habits among its inhabitants. Although some WWTPs contain a portion of industrial (YL) or slaughterhouse (SWH) wastewater, the discrepancy in microbial communities caused by this part of wastewater may not be reflected at the phylum level. To a certain extent, the microbial communities identified in effluent samples bore a striking resemblance to those found in AS samples, as bacteria present in the effluent can be viewed as the unsettled portion of AS (44, 45). Except for ST WWTP which had a temporal succession (Fig. S4), Hong Kong AS had a stable community structure. The AS of SK and SWH WWTP, as well as the AS from ST (201808 to 201901 in Fig. S4) and TP WWTP, had similar bacterial communities. The AS from YL and STL WWTP had their unique cluster patterns mightily due to the significant proportion of industrial wastewater in YL WWTP and the unique integrated fixed film activated sludge process of STL WWTP.

#### *Co-occurrence analysis reveals synchronized species-species associations in Hong Kong WWTPs*

Network analysis showed Hong Kong WWTPs had different complexity and patterns with species in ST WWTP had the most relationships while STL WWTP had the lowest one (Fig. S5; Table S5). Similar to other studies (24), positive correlations dominated relationships among species of six WWTPs while species in ST WWTP contained many negative correlations. A module in networks is a group of OTUs that have lots of connections among themselves but very few connections with other OTUs out of that group, and bacteria in the same module might carry out similar tasks (46), which means more “small

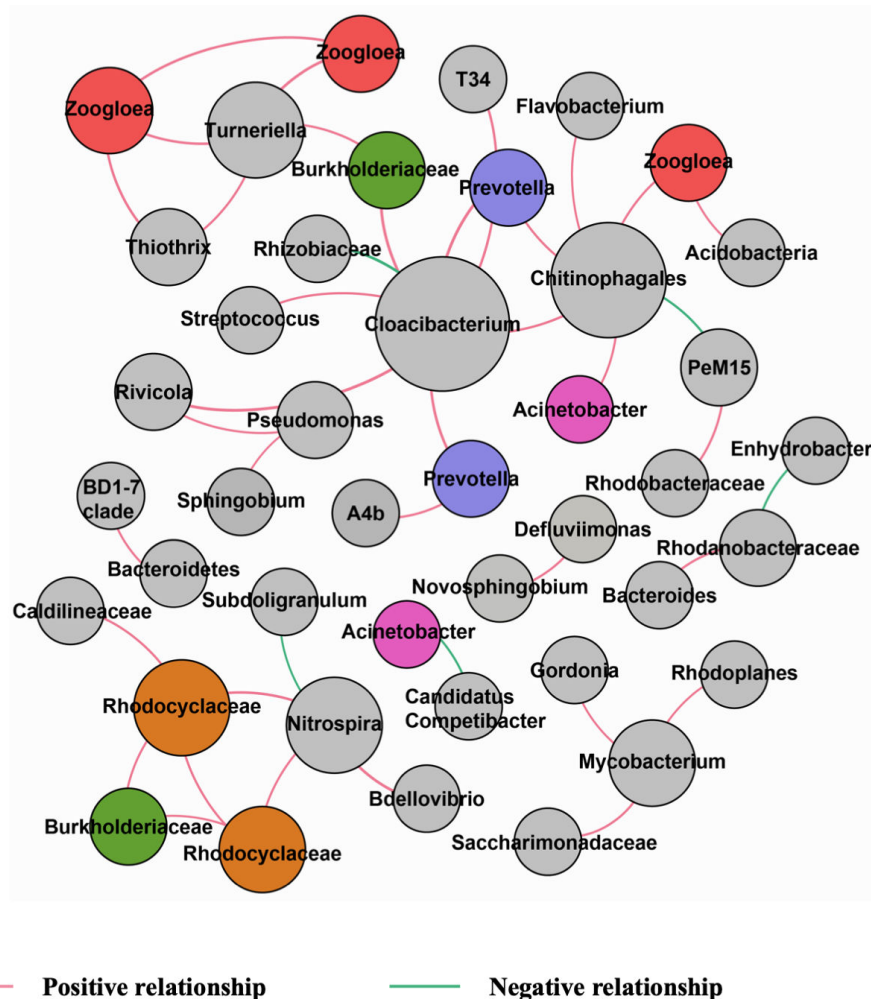




**FIG 1** Diversity analysis of six Hong Kong WWTPs. (a) The box chart of alpha diversity of influent, AS, and effluent. Outliers were marked. (b) The PCoA plot of influent, AS and effluent. The analysis was conducted by comparing pairwise microbial community dissimilarity distances using the Bray-Curtis algorithm. Each dot represented the microbial community structure of an individual sample. The colors and shapes of the dots represented WWTPs and sample types, respectively.

modules” could increase network stability. Hong Kong WWTPs had lots of modules in networks, which might contribute to the stable treatment effect.

In addition to species correlations of each WWTP, the relationships that exist among different WWTPs are more significant since they may represent solid and universal associations between two species that can greatly deepen the current understanding of microbial communities and contribute to the regulation of microbial communities. Fig. 2 shows 40 pairs of co-occurrence relationships identified by the network analysis. This co-occurrence pattern contained seven distinct groups and revealed common relationships between NOB/BFB and other bacteria in WWTPs. For instance, *Nitrospira* had occurrence patterns with two genera (*Bdellovibrio* and *Subdoligranulum*) and one core OTU belonging to the family *Rhodocyclaceae*. *Bdellovibrio* is a genus of gram-negative and obligate aerobic bacteria. One of this genus's prominent traits is that its members may feed on their hosts' biopolymers, such as proteins and nucleic acids, by preying upon other gram-negative bacteria (47). *Bdellovibrio* is putative predator of *Nitrospira* spp. (48, 49), while it could also prey on bacteria that suppress nitrification if a positive co-relationship exists between *Bdellovibrio* and *Nitrospira*. *Rhodocyclaceae* widely exists in Hong Kong WWTPs as the core bacteria with the average abundance ranging from 0.6% to 1.4%. With diverse physiological characteristics, bacteria under the *Rhodocyclaceae* family could degrade a wide range of carbon sources with electron acceptors of nitrate or nitrite and may promote nitrification to some extent. As strictly anaerobic



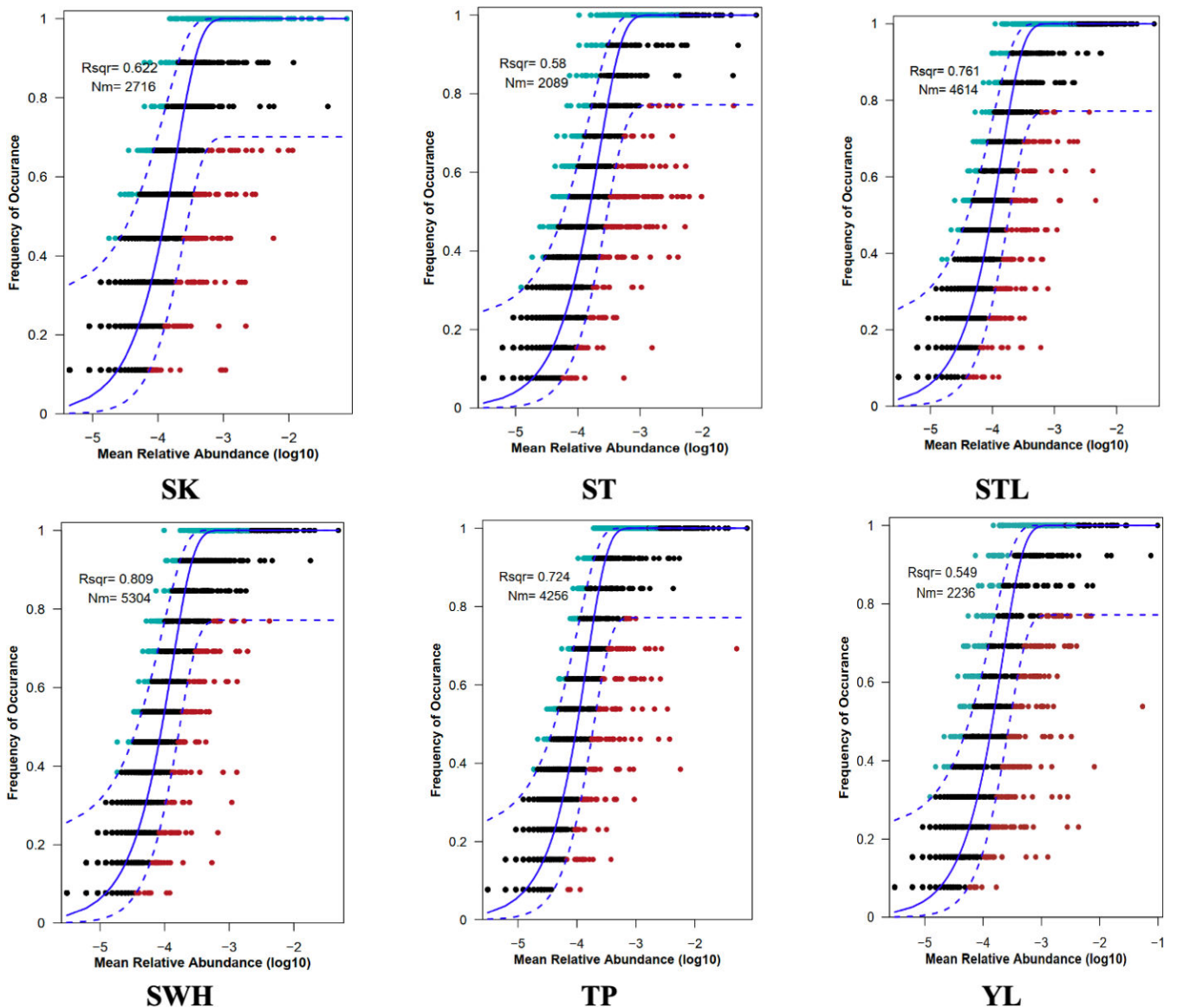
**FIG 2** Forty pairs of co-occurrence relationships identified by the network analysis. Node size was determined by relationship frequency and OTUs that occurred repeatedly were marked with the same colors. Red and green edges represented positive and negative relationships, respectively.



bacteria, *Subdoligranulum* will be suppressed by high concentrations of oxygen, while nitrification could be active under such conditions.

**Stochastic processes shape the community assembly of Hong Kong WWTPs**

Community assembly mechanism is one of the most compelling questions in the ecological study (13, 37), especially for a community that has temporal dynamics like AS. It could elucidate driving factors that shaped the community structure and provide the theoretical basis for the regulation of AS communities (50). Stochastic and deterministic processes are two key components of assembly mechanisms. Results generated from NCM (Fig. 3) showed that all six Hong Kong WWTPs were dominated by stochastic processes ( $R^2 > 0$ ). Besides, the NCM successfully estimated a large proportion of relationships between the occurrence frequency of OTUs and their relative abundance fluctuations with the explained community variance from 54.9% (YL WWTP) to 80.9% (SWH WWTP) for six Hong Kong WWTPs. The  $Nm$  had the smallest and largest values in



**FIG 3** Fit of the neutral community model (NCM) of community assembly for six Hong Kong WWTPs. The solid blue lines provided the best fit to the NCM, while the dashed blue lines showed 95% confidence intervals around the model prediction. Different colors indicated OTUs that occurred more or less frequently than expected by the NCM.  $Nm$  represents the metacommunity size times immigration and  $R^2$  was the model's fit.

ST WWTP ( $Nm = 2,089$ ) and SWH WWTP ( $Nm = 5,304$ ), respectively, indicating that the frequency of species dispersal for bacteria was lower in ST WWTP while higher in SWH WWTP.

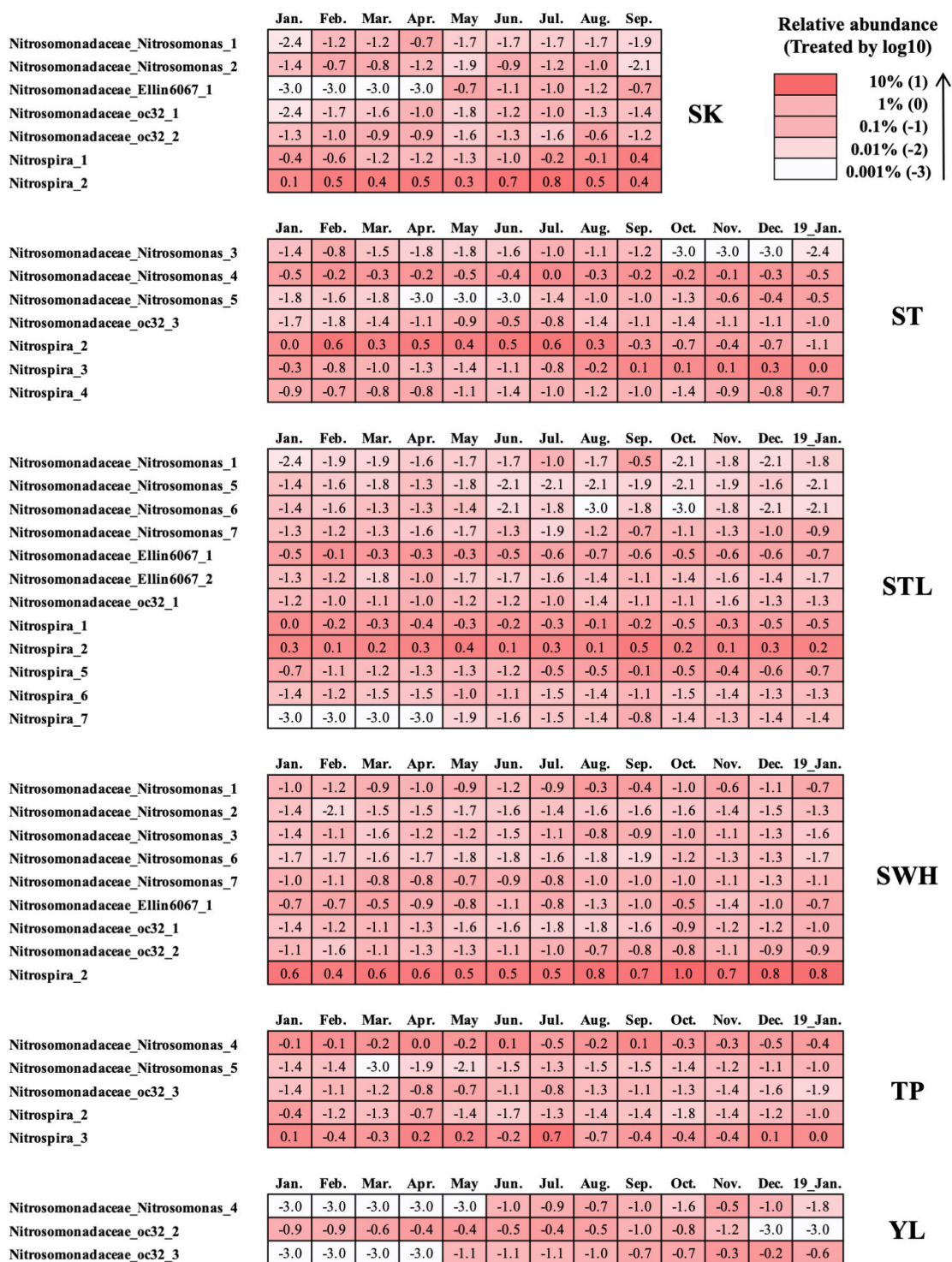
Stochastic processes dominated in communities with high diversity (51) and this stochastic dominance situation for AS was also verified in different regions (16, 24). It is reasonable that most of the OTUs in AS are dominated by stochastic processes. On the one hand, AS is an open system that receives lots of microorganisms from the influent every day and should be very vulnerable to ecological drift (52). On the other hand, functional redundancy (different populations share a similar or the same function) will increase neutrality and make functionally redundant populations more susceptible to drift from the community (13, 53). AS appears to be a microbial community that has a quite high diversity and is functionally redundant in the metabolism of lipid (40), amino acid (54), and carbohydrate metabolisms (55). Actually, previous studies often evaluated the assembly mechanism at the community level, and it would be meaningful to consider ecological processes at a lower level (e.g., the level of individual taxa or lineages) since actions of assembly processes are typically at the genotype or population level (10, 13, 18), such as the tool of Infer Community Assembly Mechanisms by Phylogenetic-bin-based null model (10). This kind of analysis may provide some novel perspectives on community assembly by finding influenced bacteria and corresponding parameters.

## The conditions of functional groups for Hong Kong and global WWTPs

### ***Abundance and distribution of AOB and NOB groups in Hong Kong and global WWTPs***

AOB and NOB are two kinds of bacteria related to nitrogen removal (56). AOB can catalyze ammonium to nitrite and NOB could oxidize the nitrite to nitrate (57). Together, they complete the whole nitrification process. In all, 39 OTUs were defined as AOB (Fig. 4) in Hong Kong WWTPs and they all belonged to the *f\_Nitrosomonadaceae* (a key member of the phylum *Proteobacteria*) with six genera: *g\_Nitrosomonas*, *g\_oc32*, *g\_Ellin6067*, *g\_DSSD61*, *g\_mle1-7*, and *g\_966-1*. Among them, *Nitrosomonas* was the most universal and typical AOB that appeared in all WWTPs, and *Nitrosomonadaceae\_Nitrosomonas\_4* in TP WWTP had the highest average abundance of 0.60%. However, the distribution of AOB across various WWTPs is highly diverse, and no unique AOB has been identified to be consistently dominant across all WWTPs. For the distribution of the AOB diversity, SWH WWTP had the most diverse AOB species across Hong Kong WWTPs in this study, although these species all had low abundance (0.07% average). However, YL WWTP had the lowest diversity with only three AOB species identified but these three AOBs had a high average abundance of 0.16%.

NOB distributes in different lineages like the genus *g\_Nitrobacter* (57), *g\_Nitrospina*, *g\_Nitrospira*, *g\_Nitrolancea*, *g\_Nitrotoga*, and *g\_Nitrococcus* (58), and has been detected in diverse environments such as soil (59), marine (60), WWTPs (61), and drinking water (62). The predominant NOB of different WWTPs may vary due to changes in wastewater types and operating parameters. For instance, a previous study showed that *Nitrotoga* was the primary NOB in Danish WWTPs (63). In another report, some full-scale WWTPs had a low abundance of both *Nitrotoga* and *Nitrospira* but *Nitrotoga* maintained high potential activities (64) and was recognized as the key nitrite oxidizer (61). The represented NOB in Hong Kong WWTPs was the Genus *Nitrospira* in Phylum *Nitrospirae*. In total, seven *Nitrospira* species were detected in Hong Kong WWTPs. Among them, *Nitrospira\_2* was the main NOB that was widely present in all WWTPs except for YL WWTP, and its abundance was up to 4.52% average in SWH WWTP. The uniqueness of SWH WWTP in AOB and NOB conditions may be due to its content of slaughterhouse wastewater. Besides, STL WWTP had the most types of NOB with five types appearing. The average abundance ratio of AOB/NOB for WWTPs in Hong Kong (except for YL WWTP) was 0.4 with the highest ratio of 1.3 for TP WWTP and the lowest ratio of 0.1 for SK WWTP. In addition to AOB and NOB, anaerobic ammonium oxidation (anammox)



**FIG 4** Composition, abundance, and distribution of AOB and NOB in Hong Kong WWTPs. The number in the figure represented the abundance and the abundance was treated by log10. Only the OTUs with an average abundance >0.01% were reserved. 19\_Jan. represented the sample collected in January 2019.

bacteria can remove ammonia from wastewater by converting nitrite and ammonium ions into diatomic nitrogen and water directly (65) to contribute to the global nitrogen cycle. Interestingly, sequencing results of the 16S rRNA gene showed that anammox bacteria only presented in YL WWTP with an extremely low abundance (0.01% average). The identified anammox bacteria belonged to the genus of *Candidatus Brocadia* in

the phylum *Planctomycetota*, which contained the first discovered anammox organism (*Candidatus Brocadia anammoxidans*).

### **AOB is more diverse than NOB for both Hong Kong and global WWTPs**

To have a deeper understanding of the nitrification process of Hong Kong and worldwide WWTPs, we examined AOB and NOB conditions in a worldwide survey for 269 WWTPs with 1,186 AS samples and then compared their results with Hong Kong WWTPs (Fig. S6; Table S6). In total, 358 AOB OTUs were defined for worldwide WWTPs with ~67% of these AOB OTUs did not have the genome information yet (Table S7). These 358 AOB OTUs are distributed in *f\_Nitrosomonadaceae*, *g\_Nitrosococcus*, *g\_Nitrospira*, and *g\_Nitrosomonas*. Besides, Hong Kong WWTPs had four unique AOB OTUs which were not detected in worldwide WWTPs (Fig. S6), although they all had low relative abundance. More interestingly, these four AOB OTUs were only detected in SWH WWTPs, indicating the special ecological niche of the AOB community in SWH WWTPs for nitrification. Moreover, it is challenging to identify a dominant AOB across WWTPs since the type and abundance of AOB exhibit a considerable amount of diversity among different WWTPs as revealed by the PCoA analysis in Fig. S7. This scenario is consistent with the results we observed in Hong Kong WWTPs.

In all, 310 NOB OTUs were detected in worldwide WWTPs, including *c\_P9 × 2b3D02*, *g\_Nitrococcus*, *g\_Nitrolancea*, *g\_Candidatus Nitrotoga*, and *g\_Nitrospira*. The *g\_Nitrospira* was identified to be the most abundant NOB in not only this study but also another research of global WWTPs (66). Among these 310 NOB OTUs, 65% of them did not have genome information yet (Table S7), indicating the importance of obtaining complete/high-quality genomes of these key species in WWTPs. Unlike AOB, NOB shows a high degree of consistency across different WWTPs with the consequence that a *Nitrospira* OTU (OTU\_6) was identified as the core taxon for worldwide WWTPs (16) by showing in 1,057 out of 1,186 WWTPs with the average and highest relative abundance of 0.9% and 7.7%, respectively. It was also the dominant NOB for Hong Kong WWTPs (*Nitrospira\_2*) with resolved genome information.

AOB appears to exhibit greater diversity and discrepancy than NOB in both Hong Kong and worldwide WWTPs: the types and quantities of AOB in different WWTPs vary greatly. Wastewater has relatively complex sources and violently fluctuating water quality with dynamic physiochemical parameters. AOB is the first executor of nitrification and high diversity could facilitate its rapid adaption to various conditions of wastewater by exploring different niches in the AS community to facilitate the nitrification process. The role of NOB is to oxidize nitrite produced by AOB, which means NOB can experience a more stable reaction environment after the performance of the AOB “buffer.” Therefore, the mainstream NOB strains are relatively uniform across different WWTPs. The condition that the AOB community was more diverse, resistant, and resilient than the NOB community was also reported in both WWTPs (67, 68) and nitrifying bioreactors (69, 70) by other researchers when they assessed the dynamic of the microbial population. In fact, the situation that the diversity of bacteria performing the first step is higher than that of the second/following step is common in the multi-step microbial reaction. For example, the methanogenic process contains four steps: hydrolysis, fermentation, syntrophic oxidation, and methanogenesis (71). Basically, the species and abundance of bacteria with the potential to proceed with each process gradually decreased and only a limited microbial group could complete the final methanogenesis step (72, 73).

### **Abundance and distribution of PAO groups in Hong Kong and global WWTPs**

Another important function of WWTPs is the removal of phosphorus, which is mainly completed by PAOs, a group of bacteria that can facilitate the removal of large amounts of phosphorus from wastewater by accumulating phosphorus within their cells as polyphosphate (19). *Candidatus Accumulibacter* (*Ca. Accumulibacter*) (74) and *Tetrasphaera* (75) are two kinds of PAOs whose phosphorus removal abilities have been identified by researchers. In summary, five *Ca. Accumulibacter* and two *Tetrasphaera* OTUs



presented in Hong Kong WWTPs and the *Ca. Accumulibacter* had a higher abundance than *Tetrasphaera* (Fig. S8). Among them, *Ca. Accumulibacter\_1* was the key PAO in Hong Kong WWTPs that showed in three WWTPs and had the highest average abundance of 1.03% in SWH WWTP. With the wider distribution, *Tetrasphaera\_1* existed in all WWTPs and showed the highest average abundance of 0.31% in ST WWTP. In all, 92 *Ca. Accumulibacter* and 33 *Tetrasphaera* were found in worldwide WWTPs with the fact that *Ca. Accumulibacter* had more diversity while *Tetrasphaera* was more prevalent (Table S8). According to the results of global data and other studies, PAOs could have >5.02% abundance in many countries that apply the enhanced biological phosphorus removal (EBPR) technology, such as Japan, Denmark (16), and Italy (75). Compared to the results of worldwide WWTPs, the abundance of PAOs was not high in Hong Kong mainly due to the lack of phosphorus removal process (i.e., EBPR) in WWTP design.

### **Abundance and distribution of BFB groups in Hong Kong and global WWTPs**

BFB groups are always present in AS and might harm wastewater treatment systems (76). Seven kinds of BFB were detected in six Hong Kong WWTPs, including *Acinetobacter* sp., *Caldilinea* sp., *Gordonia* sp., *Moraxella osloensis*, *Mycobacterium* sp., *Skermania piniformis*, and *Thiothrix* sp. (Fig. S9). The most universal BFBs in Hong Kong WWTPs were *Acinetobacter* sp. and *Mycobacterium* sp., and these two BFBs were persistent in AS and detected in all WWTPs. The highest average abundance of *Acinetobacter* sp. and *Mycobacterium* sp. both showed in STL WWTP with an abundance of 1.22% and 0.90%, respectively. Among six Hong Kong WWTPs, STL WWTP suffered the most serious BFB problem with the highest type (five types) and abundance (0.34% average) of BFB, while SWH and YL WWTPs were much lower influenced by BFB compared with other WWTPs. Another important and common BFB that is worthy of attention is *Gordonia* sp. We detected *Gordonia* sp. in four Hong Kong WWTPs (SK, ST, STL, and TP) and further results showed *Gordonia* sp. seemed to have seasonal dynamic changes, that is, it had higher abundance in winter-spring while lower abundance in summer-autumn. This situation was also observed in a 5-year BFB survey of ST WWTP (77).

In total, 29 genera BFB (Table S9) were detected in global WWTPs with the top three genera of *Acinetobacter* (170 OTUs), *Deftluviococcus* (72 OTUs), and *Thiothrix* (63 OTUs). *Acinetobacter* was the most diverse BFB for both Hong Kong and global WWTPs, which showed in almost all WWTPs (99.3%) and had an average abundance of 0.22%. In addition to being a notorious foaming bacterium, *Acinetobacter* is a potential environmental pollutant since it contains antibiotic resistance genes and can be pathogens (78). Compared to filamentous bacteria, it is more common to study *Deftluviococcus* as the glycogen accumulating organisms, which shows in not only EBPR communities but also WWTPs operated at high temperatures or with high industrial loads (79). *Thiothrix* is also a common BFB and is detected in global WWTPs with an abundance of up to 41.3% in Santiago, Chile. Actually, the abundance of some BFB genera could be underestimated due to V4 primer bias (66), such as *Leptothrix* and *Sphaerotilus*. Interestingly, researchers reported that the growth of some BFB was related to particular treatment processes of wastewater. For example, *Ca. Microthrix* obtains good growth under microaerophilic conditions and rarely shows in WWTPs that are designed to remove carbon only (80, 81).

In fact, all the above discussions are based on the relative abundance of the 16S rRNA gene. It is more meaningful to count functional bacteria at the cell level with absolute quantification, especially for an engineering system like AS. The 16S rRNA gene has multiple copies in cells and amplification bias generated by PCR may mislead our judgment of ecological questions by providing distorted quantitative relationships. Although many tools like rrnDB (82) are available to conduct the correction, we have to be careful about the resulting false positives when trying to convert 16S rRNA gene numbers to cell numbers (83). Absolute quantification has critical significance in microbial research when researchers want to compare samples collected from different sites like this study or long-term studies. The enrichment of taxa (increase in relative abundance) does not necessarily relate to the outgrowth of taxa (increase in absolute

abundance) (84), and thus direct comparisons of relative abundance for different samples may lead to wrong conclusions.

### The MDM conditions for Hong Kong and global WWTPs

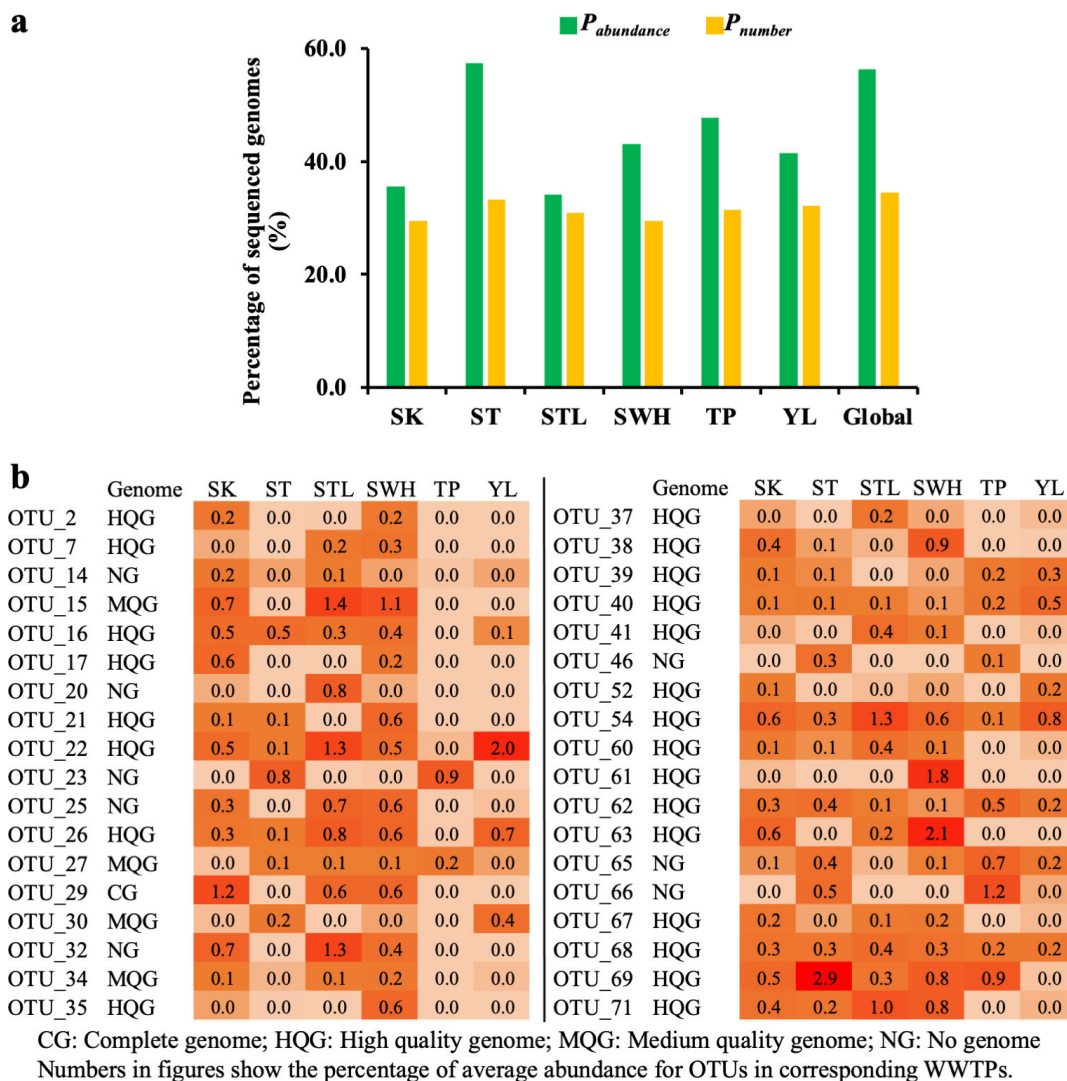
Our previous research defined the “wanted list” of global WWTPs and revealed that the median proportions of the genome-sequenced cells and taxa (100% identity and 100% coverage in 16S rRNA gene region) in global AS reached 56.3% and 34.5%, respectively (23). Fig. 5a shows ST WWTP had the highest genome-sequenced proportions among six Hong Kong WWTPs since the deep sequencing (up to 500 Gb) on ST AS samples had been performed to obtain numerous high-quality and complete genomes by the hybrid assembly of short/long reads (85). The genome-sequenced proportions of TP and YL WWTPs were also high since they had similar bacterial flora structures with ST WWTP. The remaining three WWTPs of SK, STL, and SWH had lower genome-sequenced proportions than global WWTPs (23), indicating the fact that the great majority of species in Hong Kong WWTPs lack physiological characteristics and only a few microorganisms could be characterized based on available genomes and pure cultures. For the distribution of the global “wanted list” in Hong Kong WWTPs, 36 out of 71 OTUs that do not have isolates yet could be found with >0.1% average relative abundance (Fig. 5b; Table S10). For instance, OTU\_69 and OTU\_63 had an average relative abundance of 2.9% in ST WWTP and 2.1% in SWH WWTP, respectively. With such a high abundance, Hong Kong WWTPs could make great contributions to the genome mining of microbiota in AS by both sequencing and isolation (86), especially for the eight OTUs that do not have any genome information in the present database.

Although, already, there have been lots of microbiome surveys in WWTPs, the importance of sub-genus level diversity in the AS ecosystem is largely unknown. Since functional traits are often conserved at low taxonomic ranks (87), knowledge about the role of individual species in the ecosystem is of crucial importance to understanding the ecosystem function, resilience to disturbances, and biological nutrient transformation routes. The significance of species-level microbiome composition was emphasized by recent studies. For instance, *comammox* (88) and *Tetrasphaera* (89) contain certain crucial important functions to the ecosystem, and these functions are only exhibited by some members of the genus. Partial regions of the 16S rRNA gene provide poor resolution at the species level, while scientists have taken note of this and have built many databases such as GTDB (35), the ribosomal RNA operons (90), and the full-length 16S rRNA gene (91).

### The effects between operational parameters and OTUs' abundance for Hong Kong and global WWTPs

The RDA was conducted between functional OTUs and operational parameters for Hong Kong WWTPs to highlight the top three important operational parameters: temperature, mean cell residence time, and hydraulic retention time (Fig. S10). Figure S10 shows the temperature could affect the abundance of *Tetrasphaera*, *Nitrospira*, and *Gordonia* in six Hong Kong WWTPs. To further verify the relationships between bacteria and operational parameters with long-term data, we extracted the metagenomic sequencing data of ST AS (over 9 years) from the former research (40) to show the correlation of temperature and abundance with three MAGs (Fig. 6; Table S11). The situation that *Tetrasphaera* (Fig. 6a) has a negative relationship with temperature has been widely observed and reported in global WWTPs (92, 93). It seems that a temperate environment (10°C–20°C) probably supports the proliferation of *Tetrasphaera*, but *Tetrasphaera* may fail the growing competition with other microbes under high temperatures (75). According to Fig. 6b, the abundance of *Nitrospira* had a positive relationship with temperature, which indicated *Nitrospira* could have a better performance under high temperatures and provided a possible explanation for nitrification failure in winter with solid evidence (24). Different from *Nitrospira*, *Gordonia* (Fig. 6c) had a high abundance under low temperatures, which

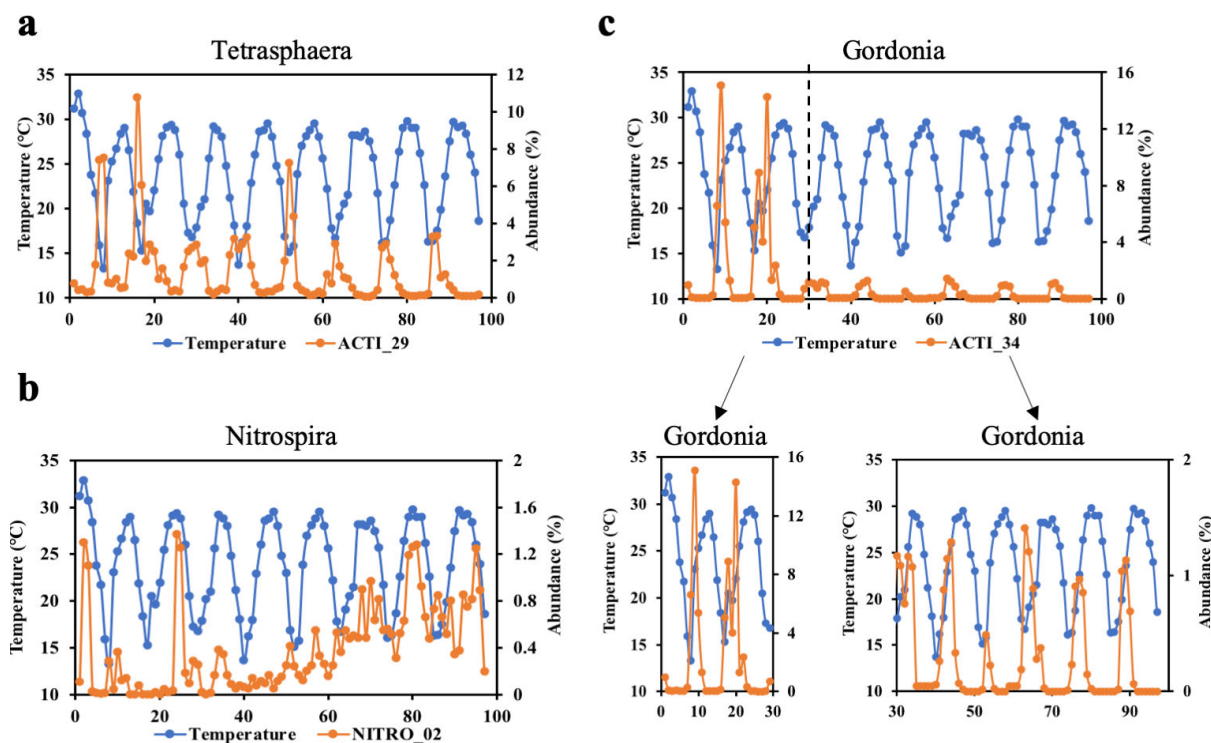




**FIG 5** MDM conditions of six Hong Kong WWTPs. (a)  $P_{number}$  and  $P_{abundance}$  results of Hong Kong and global WWTPs.  $P_{number}$  and  $P_{abundance}$  were used to represent the genome-sequenced proportions of taxa (total OTU number) number and corresponding cell number (total relative abundance sum) sum. (b) Distribution of 36 OTUs from global “wanted list” in Hong Kong WWTPs.

was consistent with our observation that sludge foaming had a high frequency in winter for Hong Kong (1) and other WWTPs (94).

Actually, the abundance change in the microbial community is the interaction consequence of many operational parameters and it is particularly difficult to elaborate on the influence of one operational parameter specifically. More long-term data on global WWTPs are needed to track continuous dynamic changes of specific bacteria in one WWTP and verify stable relationships in multiple WWTPs with different treatment processes and operational parameters. RDA in this study could provide a solution to find potential crucial parameters for technical improvements and more data analysis tools should be applied to reveal key connections among the community in the big data era. For instance, researchers tried to predict the community composition with machine/deep learning models trained by environmental parameters (27, 95, 96). All this valuable information can help engineers understand microbial relationships deeply and control the performance of WWTPs better.



**FIG 6** Relationships of temperature and abundance of MAGs in ST AS with over 9-year data. Three MAGs were included: (a) *Tetrasphaera*, (b) *Nitrospira*, and (c) *Gordonia*.

## Conclusions

A 13-month bacterial community survey of six full-scale WWTPs in Hong Kong was conducted and results were compared to data from worldwide WWTPs and the long-term survey of ST WWTP to get the following conclusions.

- The AS bacterial community structures of six Hong Kong WWTPs were markedly distinct from one another.
- Forty pairs of co-occurrence relationships identified in multiple WWTPs revealed widespread microbial associations in AS.
- Stochastic processes took large proportions for the bacterial community assembly.
- The abundance and distribution of the functional bacteria in worldwide and Hong Kong WWTPs were different. The AOB was more diverse than NOB for both Hong Kong and worldwide WWTPs. This fact may reveal a fundamental truth that the functional bacteria responsible for downstream reactions in multi-step biochemical processes are more conserved.
- Of 71 OTUs in the global “wanted list,” 36 did not have isolates yet but had >0.1% average abundance in Hong Kong WWTPs.
- Operational parameters had important effects on OTUs’ abundance for both Hong Kong and worldwide WWTPs, such as the temperature to the genera of *Tetrasphaera*, *Gordonia* and *Nitrospira*.
- Further research could reveal novel insights into bacterial communities in WWTPs with tools of community assembly pipelines at the lineage level, absolute quantification and machine learning.

## ACKNOWLEDGMENTS

The authors sincerely thank Miss Vicky Fung, Dr. Lei Liu, Dr. Yu Yang, and other labmates for their technical support throughout the whole project.

Shuxian Li would like to thank The University of Hong Kong for the Postgraduate Studentship. Yulin Zhang wishes to thank The University of Hong Kong for the Postdoctoral fellowship. This study was financially supported by the Hong Kong Theme Based Research Scheme (T21-604/19-R), the Hong Kong General Research Fund (GRF17206120), and the Hong Kong DSD (DEMP-2017-03).

T.Z. and Y.D. designed the project. Y.D. and Y.Z. performed the experiments. Y.Z., C.W., and S.L. analyzed the results and wrote the manuscript. F.T.K.L. arranged sampling processes and provided data operational parameters data of WWTPs. Jizhong Zhou revised the manuscript. The authors read and approved the final manuscript.

## AUTHOR AFFILIATIONS

<sup>1</sup>Environmental Microbiome Engineering and Biotechnology Lab, Department of Civil Engineering, The University of Hong Kong, Hong Kong, China

<sup>2</sup>Drainage Services Department, The Government of the Hong Kong Special Administrative Region of the People's Republic of China, Wanchai, Hong Kong, China

<sup>3</sup>Institute for Environmental Genomics, Department of Microbiology and Plant Biology, and School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, Oklahoma, USA

<sup>4</sup>Macau Institute for Applied Research in Medicine and Health, Macau University of Science and Technology, Macau, China

## AUTHOR ORCIDs

Yulin Zhang  <http://orcid.org/0000-0003-3476-8811>

Tong Zhang  <http://orcid.org/0000-0003-1148-4322>

## FUNDING

Funder	Grant(s)	Author(s)
<a href="#">Hong Kong Theme Based Research Scheme</a>	T21-604/19-R	Tong Zhang
<a href="#">Hong Kong General Research Fund</a>	GRF17206120	Tong Zhang
<a href="#">Hong Kong DSD</a>	DEMP-2017-03	Tong Zhang

## AUTHOR CONTRIBUTIONS

Yulin Zhang, Data curation, Methodology, Visualization, Writing – original draft, Writing – review and editing | Yu Deng, Data curation, Methodology, Project administration | Chunxiao Wang, Methodology, Resources, Writing – review and editing | Shuxian Li, Data curation, Methodology, Visualization, Writing – review and editing | Frankie T. K. Lau, Methodology, Resources, Writing – review and editing | Jizhong Zhou, Data curation, Methodology, Writing – original draft, Writing – review and editing | Tong Zhang, Methodology, Visualization, Writing – review and editing

## DATA AVAILABILITY

The raw data of 16S rRNA reads could be downloaded from NCBI with the BioProject ID [PRJNA1012295](#).

## ADDITIONAL FILES

The following material is available [online](#).

### Supplemental Material

**Supplemental material** ([mSystems01333-23-s0001.pdf](#)). Supplemental text and Figures S1-S13.

## Supplemental tables (mSystems01333-23-s0002.xlsx). Tables S1–S11.

## REFERENCES

- Jiang X-T, Ye L, Ju F, Wang Y-L, Zhang T. 2018. Toward an intensive longitudinal understanding of activated sludge bacterial assembly and dynamics. *Environ Sci Technol* 52:8224–8232. <https://doi.org/10.1021/acs.est.7b05579>
- Ju F, Zhang T. 2015. Bacterial assembly and temporal dynamics in activated sludge of a full-scale municipal wastewater treatment plant. *ISME J* 9:683–695. <https://doi.org/10.1038/ismej.2014.162>
- Wang Y, Niu Q, Zhang X, Liu L, Wang Y, Chen Y, Negi M, Figeys D, Li Y-Y, Zhang T. 2019. Exploring the effects of operational mode and microbial interactions on bacterial community assembly in a one-stage partial-nitritation Anammox reactor using integrated multi-Omics. *Microbiome* 7:122. <https://doi.org/10.1186/s40168-019-0730-6>
- Yu X, Polz MF, Alm EJ. 2019. Interactions in self-assembled microbial communities saturate with diversity. *ISME J* 13:1602–1617. <https://doi.org/10.1038/s41396-019-0356-5>
- Zhang B, Ning D, Yang Y, Van Nostrand JD, Zhou J, Wen X. 2020. Biodegradability of wastewater determines microbial assembly mechanisms in full-scale wastewater treatment plants. *Water Res* 169:115276. <https://doi.org/10.1016/j.watres.2019.115276>
- Ibarbalz FM, Figuerola ELM, Erijman L. 2013. Industrial activated sludge exhibit unique bacterial community composition at high taxonomic ranks. *Water Res* 47:3854–3864. <https://doi.org/10.1016/j.watres.2013.04.010>
- Barberán A, Bates ST, Casamayor EO, Fierer N. 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J* 6:343–351. <https://doi.org/10.1038/ismej.2011.119>
- Ruan Q, Dutta D, Schwalbach MS, Steele JA, Fuhrman JA, Sun F. 2006. Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* 22:2532–2538. <https://doi.org/10.1093/bioinformatics/btl417>
- Zieliński W, Korzeniewska E, Harnisz M, Drzymała J, Felis E, Bajkacz S. 2021. Wastewater treatment plants as a reservoir of Integrase and antibiotic resistance genes - an epidemiological threat to workers and environment. *Environ Int* 156:106641. <https://doi.org/10.1016/j.envint.2021.106641>
- Ning D, Yuan M, Wu L, Zhang Y, Guo X, Zhou X, Yang Y, Arkin AP, Firestone MK, Zhou J. 2020. A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat Commun* 11:4717. <https://doi.org/10.1038/s41467-020-18560-z>
- Assess HA, Selvarajan R, Nyoni H, Ntushelo K, Mamba BB, Msagati TAM. 2019. Diversity, co-occurrence and implications of fungal communities in wastewater treatment plants. *Sci Rep* 9, 14056. <https://doi.org/10.1038/s41598-019-50624-z>
- Pholchan MK, Baptista J de C, Davenport RJ, Sloan WT, Curtis TP. 2013. Microbial community assembly, theory and rare functions. *Front Microbiol* 4:68. <https://doi.org/10.3389/fmicb.2013.00068>
- Zhou J, Ning D. 2017. Stochastic community assembly: does it matter in microbial ecology? *Microbiol Mol Biol Rev* 81:e00002-17. <https://doi.org/10.1128/MMBR.00002-17>
- Johnston J, LaPara T, Behrens S. 2019. Composition and dynamics of the activated sludge microbiome during seasonal nitrification failure. *Sci Rep* 9, 4565. <https://doi.org/10.1038/s41598-019-40872-4>
- Griffin JS, Wells GF. 2017. Regional synchrony in full-scale activated sludge bioreactors due to deterministic microbial community assembly. *ISME J* 11:500–511. <https://doi.org/10.1038/ismej.2016.121>
- Wu L, Ning D, Zhang B, Li Y, Zhang P, Shan X, Zhang Q, Brown MR, Li Z, Van Nostrand JD, et al. 2019. Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nat Microbiol* 4:1183–1195. <https://doi.org/10.1038/s41564-019-0426-5>
- Zhang B, Ning D, Van Nostrand JD, Sun C, Yang Y, Zhou J, Wen X. 2020. Biogeography and assembly of microbial communities in wastewater treatment plants in China. *Environ Sci Technol* 54:5884–5892. <https://doi.org/10.1021/acs.est.9b07950>
- Ofiteru ID, Lunn M, Curtis TP, Wells GF, Criddle CS, Francis CA, Sloan WT. 2010. Combined niche and neutral effects in a microbial wastewater treatment community. *Proc Natl Acad Sci U S A* 107:15345–15350. <https://doi.org/10.1073/pnas.1000604107>
- Nielsen PH, Saunders AM, Hansen AA, Larsen P, Nielsen JL. 2012. Microbial communities involved in enhanced biological phosphorus removal from wastewater—a model system in environmental biotechnology. *Curr Opin Biotechnol* 23:452–459. <https://doi.org/10.1016/j.copbio.2011.11.027>
- Parks DH, Rinke C, Chuvpochina M, Chaumeil P-A, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW. 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol* 2:1533–1542. <https://doi.org/10.1038/s41564-017-0012-7>
- Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu W-T, Eisen JA, Hallam SJ, Kyrpidis NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499:431–437. <https://doi.org/10.1038/nature12352>
- Zamkovaya T, Foster JS, de Crécy-Lagard V, Conesa A. 2021. A network approach to elucidate and prioritize microbial dark matter in microbial communities. *ISME J* 15:228–244. <https://doi.org/10.1038/s41396-020-00777-x>
- Zhang Y, Wang Y, Tang M, Zhou J, Zhang T. 2023. The microbial dark matter and “wanted list” in worldwide wastewater treatment plants. *Microbiome* 11:59. <https://doi.org/10.1186/s40168-023-01503-3>
- Sun C, Zhang B, Ning D, Zhang Y, Dai T, Wu L, Li T, Liu W, Zhou J, Wen X. 2021. Seasonal dynamics of the microbial community in two full-scale wastewater treatment plants: diversity, composition, phylogenetic group based assembly and co-occurrence pattern. *Water Res* 200:117295. <https://doi.org/10.1016/j.watres.2021.117295>
- Nielsen PH, Mielczarek AT, Kragelund C, Nielsen JL, Saunders AM, Kong Y, Hansen AA, Vollertsen J. 2010. A conceptual ecosystem model of microbial communities in enhanced biological phosphorus removal plants. *Water Res* 44:5070–5088. <https://doi.org/10.1016/j.watres.2010.07.036>
- Chen Y, Lan S, Wang L, Dong S, Zhou H, Tan Z, Li X. 2017. A review: driving factors and regulation strategies of microbial community structure and dynamics in wastewater treatment systems. *Chemosphere* 174:173–182. <https://doi.org/10.1016/j.chemosphere.2017.01.129>
- Zhong S, Zhang K, Bagheri M, Burken JG, Gu A, Li B, Ma X, Marrone BL, Ren ZJ, Schrier J, Shi W, Tan H, Wang T, Wang X, Wong BM, Xiao X, Yu X, Zhu J-J, Zhang H. 2021. Machine learning: new ideas and tools in environmental science and engineering. *Environ Sci Technol* 55:12741–12754. <https://doi.org/10.1021/acs.est.1c01339>
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583. <https://doi.org/10.1038/nmeth.3869>
- Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37:852–857. <https://doi.org/10.1038/s41587-019-0209-9>
- Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. <https://doi.org/10.7717/peerj.2584>
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>
- Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. 2019. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res* 47:W5–W10. <https://doi.org/10.1093/nar/gkz342>
- Price MN, Dehal PS, Arkin AP. 2010. Fasttree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>



34. Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49:W293–W296. <https://doi.org/10.1093/nar/gkab301>
35. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. 2020. A complete domain-to-species taxonomy for bacteria and archaea. *Nat Biotechnol* 38:1079–1086. <https://doi.org/10.1038/s41587-020-0501-8>
36. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>
37. Chen W, Ren K, Isabwe A, Chen H, Liu M, Yang J. 2019. Stochastic processes shape microeukaryotic community assembly in a subtropical river across wet and dry seasons. *Microbiome* 7:138. <https://doi.org/10.1186/s40168-019-0749-8>
38. Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. *ICWSM* 3:361–362. <https://doi.org/10.1609/icwsm.v3i1.13937>
39. Kurtz ZD, Müller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA. 2015. Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol* 11:e1004226. <https://doi.org/10.1371/journal.pcbi.1004226>
40. Wang Y, Ye J, Ju F, Liu L, Boyd JA, Deng Y, Parks DH, Jiang X, Yin X, Woodcroft BJ, Tyson GW, Hugenholtz P, Polz MF, Zhang T. 2021. Successional dynamics and alternative stable states in a saline activated sludge microbial community over 9 years. *Microbiome* 9:199. <https://doi.org/10.1186/s40168-021-01151-5>
41. Newton RJ, McLellan SL, Dila DK, Vineis JH, Morrison HG, Eren AM, Sogin ML. 2015. Sewage reflects the microbiomes of human populations. *mBio* 6:e02574. <https://doi.org/10.1128/mBio.02574-14>
42. Su J-Q, An X-L, Li B, Chen Q-L, Gillings MR, Chen H, Zhang T, Zhu Y-G. 2017. Metagenomics of urban sewage identifies an extensively shared antibiotic resistome in China. *Microbiome* 5:84. <https://doi.org/10.1186/s40168-017-0298-y>
43. LaMartina EL, Mohaimani AA, Newton RJ. 2021. Urban wastewater bacterial communities assemble into seasonal steady states. *Microbiome* 9:116. <https://doi.org/10.1186/s40168-021-01038-5>
44. Kristensen JM, Nierychlo M, Albertsen M, Nielsen PH. 2020. Bacteria from the genus *Arcobacter* are abundant in effluent from wastewater treatment plants. *Appl Environ Microbiol* 86:e03044-19. <https://doi.org/10.1128/AEM.03044-19>
45. Ye L, Mei R, Liu W-T, Ren H, Zhang X-X. 2020. Machine learning-aided analyses of thousands of draft genomes reveal specific features of activated sludge processes. *Microbiome* 8:16. <https://doi.org/10.1186/s40168-020-0794-3>
46. Deng Y, Jiang Y-H, Yang Y, He Z, Luo F, Zhou J. 2012. Molecular ecological network analyses. *BMC Bioinformatics* 13:113. <https://doi.org/10.1186/1471-2105-13-113>
47. Koval SF, Hynes SH, Flannagan RS, Pasternak Z, Davidov Y, Jurkevitch E. 2013. *Bdellovibrio exovorax* sp. nov., a novel predator of *Caulobacter crescentus*. *Int J Syst Evol Microbiol* 63:146–151. <https://doi.org/10.1099/ijs.0.039701-0>
48. Cohen Y, Pasternak Z, Müller S, Hübschmann T, Schattenberg F, Sivakala KK, Abed-Rabbo A, Chatzinotas A, Jurkevitch E. 2021. Community and single cell analyses reveal complex predatory interactions between bacteria in high diversity systems. *Nat Commun* 12:5481. <https://doi.org/10.1038/s41467-021-25824-9>
49. Dolinšek J, Lagkouvardos I, Wanek W, Wagner M, Daims H. 2013. Interactions of nitrifying bacteria and heterotrophs: identification of a *Micavibrio*-like putative predator of *Nitrospira* spp. *Appl Environ Microbiol* 79:2027–2037. <https://doi.org/10.1128/AEM.03408-12>
50. Isazadeh S, Jauffur S, Frigon D. 2016. Bacterial community assembly in activated sludge: mapping beta diversity across environmental variables. *Microbiologyopen* 5:1050–1060. <https://doi.org/10.1002/mbo3.388>
51. Xun W, Li W, Xiong W, Ren Y, Liu Y, Miao Y, Xu Z, Zhang N, Shen Q, Zhang R. 2019. Diversity-triggered deterministic bacterial assembly constrains community functions. *Nat Commun* 10:3833. <https://doi.org/10.1038/s41467-019-11787-5>
52. Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S. 2013. Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* 77:342–356. <https://doi.org/10.1128/MMBR.00051-12>
53. Louca S, Polz MF, Mazel F, Albright MB, Huber JA, O'Connor MI, Ackermann M, Hahn AS, Srivastava DS, Crowe SA, Doebeli M, Parfrey LW. 2018. Function and functional redundancy in microbial systems. *Nat Ecol Evol* 2:936–943. <https://doi.org/10.1038/s41559-018-0519-1>
54. Chen H, Wang M, Chang S. 2020. Disentangling community structure of ecological system in activated sludge: core communities, functionality, and functional redundancy. *Microb Ecol* 80:296–308. <https://doi.org/10.1007/s00248-020-01492-y>
55. Delgado - Baquerizo M, Giaramida L, Reich PB, Khachane AN, Hamonts K, Edwards C, Lawton LA, Singh BK, Brophy C. 2016. Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. *J Ecol* 104:936–946. <https://doi.org/10.1111/1365-2745.12585>
56. Kuypers MMM, Marchant HK, Kartal B. 2018. The microbial nitrogen-cycling network. *Nat Rev Microbiol* 16:263–276. <https://doi.org/10.1038/nrmicro.2018.9>
57. Grunditz C, Dalhammar G. 2001. Development of nitrification inhibition assays using pure cultures of *Nitrosomonas* and *Nitrobacter*. *Water Res* 35:433–440. [https://doi.org/10.1016/S0043-1354\(00\)00312-2](https://doi.org/10.1016/S0043-1354(00)00312-2)
58. Daims H, Lückner S, Wagner M. 2016. A new perspective on microbes formerly known as nitrite-oxidizing bacteria. *Trends Microbiol* 24:699–712. <https://doi.org/10.1016/j.tim.2016.05.004>
59. Xia W, Zhang C, Zeng X, Feng Y, Weng J, Lin X, Zhu J, Xiong Z, Xu J, Cai Z, Jia Z. 2011. Autotrophic growth of nitrifying community in an agricultural soil. *ISME J* 5:1226–1236. <https://doi.org/10.1038/ismej.2011.5>
60. Füssel J, Lückner S, Yilmaz P, Nowka B, van Kessel M, Bourceau P, Hach PF, Littmann S, Berg J, Spieck E, Daims H, Kuypers MMM, Lam P. 2017. Adaptability as the key to success for the ubiquitous marine nitrite oxidizer *Nitroccoccus*. *Sci Adv* 3:e1700807. <https://doi.org/10.1126/sciadv.1700807>
61. Lückner S, Schwarz J, Gruber-Dörninger C, Spieck E, Wagner M, Daims H. 2015. *Nitrotoga*-like bacteria are previously unrecognized key nitrite oxidizers in full-scale wastewater treatment plants. *ISME J* 9:708–720. <https://doi.org/10.1038/ismej.2014.158>
62. Tatari K, Musovic S, Gülay A, Dechesne A, Albrechtsen H-J, Smets BF. 2017. Density and distribution of nitrifying guilds in rapid sand filters for drinking water production: dominance of *Nitrospira* spp. *Water Res* 127:239–248. <https://doi.org/10.1016/j.watres.2017.10.023>
63. Saunders AM, Albertsen M, Vollertsen J, Nielsen PH. 2016. The activated sludge ecosystem contains a core community of abundant organisms. *ISME J* 10:11–20. <https://doi.org/10.1038/ismej.2015.117>
64. Lawson CE, Strachan BJ, Hanson NW, Hahn AS, Hall ER, Rabinowitz B, Mavinic DS, Ramey WD, Hallam SJ. 2015. Rare taxa have potential to make metabolic contributions in enhanced biological phosphorus removal ecosystems. *Environ Microbiol* 17:4979–4993. <https://doi.org/10.1111/1462-2920.12875>
65. Kuenen JG. 2008. Anammox bacteria: from discovery to application. *Nat Rev Microbiol* 6:320–326. <https://doi.org/10.1038/nrmicro1857>
66. Dueholm MKD, Nierychlo M, Andersen KS, Rudkjøbing V, Knutsson S, Albertsen M, Nielsen PH, MiDAS Global Consortium. 2022. MiDAS 4: a global catalogue of full-length 16S rRNA gene sequences and taxonomy for studies of bacterial communities in wastewater treatment plants. *Nat Commun* 13:1908. <https://doi.org/10.1038/s41467-022-29438-7>
67. Cotto I, Vilardi KJ, Huo L, Fogarty EC, Khunjar W, Wilson C, De Clippeleir H, Gilmore K, Bailey E, Lückner S, Pinto AJ. 2023. Low diversity and microdiversity of *comammox* bacteria in wastewater systems suggest specific adaptations within the *Ca. Nitrospira nitrosa* cluster. *Water Res* 229:119497. <https://doi.org/10.1016/j.watres.2022.119497>
68. Zhang B, Sun C, Lin H, Liu W, Qin W, Chen T, Yang T, Wen X. 2023. Differences in distributions, assembly mechanisms, and putative interactions of AOB and NOB at a large spatial scale. *Front Environ Sci* 17:122. <https://doi.org/10.1007/s11783-023-1722-0>
69. Bejarano Ortiz DI, Martínez Jardines MA, Cuervo López F de M, Texier A-C. 2020. Biological ammonium and sulfide oxidation in a nitrifying sequencing batch reactor: kinetic and microbial population dynamics assessments. *Chemosphere* 253:126637. <https://doi.org/10.1016/j.chemosphere.2020.126637>

70. Santillan E, Phua WX, Constancias F, Wuertz S. 2020. Sustained organic loading disturbance favors nitrite accumulation in bioreactors with variable resistance, recovery and resilience of nitrification and nitrifiers. *Sci Rep* 10:21388. <https://doi.org/10.1038/s41598-020-78314-1>
71. Vanwonterghem I, Jensen PD, Rabaey K, Tyson GW. 2016. Genome-centric resolution of microbial diversity, metabolism and interactions in anaerobic digestion. *Environ Microbiol* 18:3144–3158. <https://doi.org/10.1111/1462-2920.13382>
72. Wang C, Wang Y, Wang Y, Cheung K-K, Ju F, Xia Y, Zhang T. 2020. Genome-centric microbiome analysis reveals solid retention time (SRT)-shaped species interactions and niche differentiation in food waste and sludge co-digesters. *Water Res* 181:115858. <https://doi.org/10.1016/j.watres.2020.115858>
73. De Vrieze J, Verstraete W. 2016. Perspectives for microbial community composition in anaerobic digestion: from abundance and activity to connectivity. *Environ Microbiol* 18:2797–2809. <https://doi.org/10.1111/1462-2920.13437>
74. Marques R, Santos J, Nguyen H, Carvalho G, Noronha JP, Nielsen PH, Reis MAM, Oehmen A. 2017. Metabolism and ecological niche of *Tetrasphaera* and *Ca. Accumulibacter* in enhanced biological phosphorus removal. *Water Res* 122:159–171. <https://doi.org/10.1016/j.watres.2017.04.072>
75. Liu R, Hao X, Chen Q, Li J. 2019. Research advances of *Tetrasphaera* in enhanced biological phosphorus removal: a review. *Water Res* 166:115003. <https://doi.org/10.1016/j.watres.2019.115003>
76. Guo F, Zhang T. 2012. Profiling bulking and foaming bacteria in activated sludge by high throughput sequencing. *Water Res* 46:2772–2782. <https://doi.org/10.1016/j.watres.2012.02.039>
77. Jiang X-T, Guo F, Zhang T. 2016. Population dynamics of bulking and foaming bacteria in a full-scale wastewater treatment plant over five years. *Sci Rep* 6:24180. <https://doi.org/10.1038/srep24180>
78. Pulami D, Kämpfer P, Glaeser SP. 2023. High diversity of the emerging pathogen *Acinetobacter baumannii* and other *Acinetobacter* Spp. in raw manure, biogas plants digestates, and rural and urban wastewater treatment plants with system specific antimicrobial resistance profiles. *Sci Total Environ* 859:160182. <https://doi.org/10.1016/j.scitotenv.2022.160182>
79. Maszenan AM, Bessarab I, Williams RBH, Petrovski S, Seviour RJ. 2022. The phylogeny, ecology and ecophysiology of the glycogen accumulating organism (GAO) *Deffluviococcus* in wastewater treatment plants. *Water Res* 221:118729. <https://doi.org/10.1016/j.watres.2022.118729>
80. Nierychlo M, Singleton CM, Petriglieri F, Thomsen L, Petersen JF, Peces M, Kondrotaitė Z, Dueholm MS, Nielsen PH. 2021. Low global diversity of *Candidatus Microthrix*, a troublesome filamentous organism in full-scale Wwtps. *Front Microbiol* 12:690251. <https://doi.org/10.3389/fmicb.2021.690251>
81. Rossetti S, Tomei MC, Nielsen PH, Tandoi V. 2005. “*Microthrix parvicella*”, a filamentous bacterium causing bulking and foaming in activated sludge systems: a review of current knowledge. *FEMS Microbiol Rev* 29:49–64. <https://doi.org/10.1016/j.femsre.2004.09.005>
82. Stoddard SF, Smith BJ, Hein R, Roller BRK, Schmidt TM. 2015. *rrnDB*: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res* 43:D593–D598. <https://doi.org/10.1093/nar/gku1201>
83. Louca S, Doebeli M, Parfrey LW. 2018. Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome* 6:41. <https://doi.org/10.1186/s40168-018-0420-9>
84. Props R, Kerckhof F-M, Rubbens P, De Vrieze J, Hernandez Sanabria E, Waegeman W, Monsieurs P, Hammes F, Boon N. 2017. Absolute quantification of microbial taxon abundances. *ISME J* 11:584–587. <https://doi.org/10.1038/ismej.2016.117>
85. Liu L, Wang Y, Yang Y, Wang D, Cheng SH, Zheng C, Zhang T. 2021. Charting the complexity of the activated sludge microbiome through a hybrid sequencing strategy. *Microbiome* 9:205. <https://doi.org/10.1186/s40168-021-01155-1>
86. Zhang Y, Zhang T. 2023. Culturing the uncultured microbial majority in activated sludge: a critical review. *Crit Rev Environ Sci Technol* 53:601–624. <https://doi.org/10.1080/10643389.2022.2077063>
87. Martiny AC, Treseder K, Pusch G. 2013. Phylogenetic conservatism of functional traits in microorganisms. *ISME J* 7:830–838. <https://doi.org/10.1038/ismej.2012.160>
88. Yang Y, Li M, Li H, Li X-Y, Lin J-G, Denecke M, Gu J-D. 2020. Specific and effective detection of anammox bacteria using PCR primers targeting the 16S rRNA gene and functional genes. *Sci Total Environ* 734:139387. <https://doi.org/10.1016/j.scitotenv.2020.139387>
89. Dueholm MS, Andersen KS, Mclroy SJ, Kristensen JM, Yashiro E, Karst SM, Albertsen M, Nielsen PH. 2020. Generation of comprehensive ecosystem-specific reference databases with species-level resolution by high-throughput full-length 16S rRNA gene sequencing and automated taxonomy assignment (AutoTax). *mBio* 11:e01557-20. <https://doi.org/10.1128/mBio.01557-20>
90. Kerkhof LJ, Roth PA, Deshpande SV, Bernhards RC, Liem AT, Hill JM, Häggblom MM, Webster NS, Iboronke O, Mirzoyan S, Polashock JJ, Sullivan RF. 2022. A ribosomal operon database and Megablast settings for strain-level resolution of microbiomes. *FEMS Microbes 3:xtac002*. <https://doi.org/10.1093/femsmc/xtac002>
91. Nierychlo M, Andersen KS, Xu Y, Green N, Jiang C, Albertsen M, Dueholm MS, Nielsen PH. 2020. MiDAS 3: an ecosystem-specific reference database, taxonomy and knowledge platform for activated sludge and anaerobic digesters reveals species-level microbiome composition of activated sludge. *Water Res* 182:115955. <https://doi.org/10.1016/j.watres.2020.115955>
92. Vallenet D, Calteau A, Dubois M, Amours P, Bazin A, Beuvin M, Burlot L, Bussell X, Fouteau S, Gautreau G, Lajus A, Langlois J, Planel R, Roche D, Rollin J, Rouy Z, Sabatet V, Médigue C. 2020. MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Res* 48:D579–D589. <https://doi.org/10.1093/nar/gkz926>
93. Kristiansen R, Nguyen HTT, Saunders AM, Nielsen JL, Wimmer R, Le VQ, Mclroy SJ, Petrovski S, Seviour RJ, Calteau A, Nielsen KL, Nielsen PH. 2013. A metabolic model for members of the genus *Tetrasphaera* involved in enhanced biological phosphorus removal. *ISME J* 7:543–554. <https://doi.org/10.1038/ismej.2012.136>
94. Oerther DB, de los Reyes FL, de los Reyes MF, Raskin L. 2001. Quantifying filamentous microorganisms in activated sludge before, during, and after an incident of foaming by oligonucleotide probe hybridizations and antibody staining. *Water Res* 35:3325–3336. [https://doi.org/10.1016/S0043-1354\(01\)00057-4](https://doi.org/10.1016/S0043-1354(01)00057-4)
95. Huang Y, Sheth RU, Zhao S, Cohen LA, Dabaghi K, Moody T, Sun Y, Ricaurte D, Richardson M, Velez-Cortes F, Blazejewski T, Kaufman A, Ronda C, Wang HH. 2023. High-throughput microbial culturomics using automation and machine learning. *Nat Biotechnol* 41:1424–1433. <https://doi.org/10.1038/s41587-023-01674-2>
96. Michel-Mata S, Wang X-W, Liu Y-Y, Angulo MT. 2022. Predicting microbiome compositions from species assemblages through deep learning. *iMeta* 1:e3. <https://doi.org/10.1002/imt2.3>