

OPEN

The capacity of wastewater treatment plants drives bacterial community structure and its assembly

Young Kyung Kim¹, Keunje Yoo^{1,3}, Min Sung Kim², Il Han², Minjoo Lee¹, Bo Ram Kang², Tae Kwon Lee² & Joonhong Park¹

Bacterial communities in wastewater treatment plants (WWTPs) affect plant functionality through their role in the removal of pollutants from wastewater. Bacterial communities vary extensively based on plant operating conditions and influent characteristics. The capacity of WWTPs can also affect the bacterial community via variations in the organic or nutrient composition of the influent. Despite the importance considering capacity, the characteristics that control bacterial community assembly are largely unknown. In this study, we discovered that bacterial communities in WWTPs in Korea and Vietnam, which differ remarkably in capacity, exhibit unique structures and interactions that are governed mainly by the capacity of WWTPs. Bacterial communities were analysed using 16S rRNA gene sequencing and exhibited clear differences between the two regions, with these differences being most pronounced in activated sludge. We found that capacity contributed the most to bacterial interactions and community structure, whereas other factors had less impact. Co-occurrence network analysis showed that microorganisms from high-capacity WWTPs are more interrelated than those from low-capacity WWTPs, which corresponds to the tighter clustering of bacterial communities in Korea. These results will contribute to the understanding of bacterial community assembly in activated sludge processing.

Wastewater treatment encompasses the processes that convert contaminated water into a sufficiently clean state that can be discharged to surface water with minimal adverse environmental impact. Wastewater treatment plants (WWTPs) employ a series of processes, namely, preliminary treatment, primary settling, secondary treatment, secondary clarification, tertiary treatment, disinfection, and sludge processing, each of which is designed to remove different types of pollutants. Of these processes, secondary and tertiary treatments are where most organic matter, nutrients, and other pollutants are removed, and these are commonly achieved using biological methods^{1,2}. Thus, bacterial community composition, diversity, and dynamics, which are shaped by both operating conditions and influent characteristics^{3,4}, are the major factors that determine the performance of a wide range of biological treatments and activated sludge (AS) processing^{5–8}. Many studies have analyzed the responses of the key microorganisms for wastewater treatment under different WWTP operation practices^{2,9}.

Capacity is a complex function of physical constraints (e.g. reactor volumes and equipment capacities), influent characteristics (e.g. microbial composition and organic type) and operation factors (e.g. sludge retention time [SRT] and hydraulic retention times [HRT]) that are strongly associated with sludge processes including sludge flocculation, process stability, and microbial flexibility^{10,11}. WWTP capacities around the world are inherently variable. In Seoul, which has a population of 10 million, four extremely large-scale wastewater treatment facilities have been operated. Actually, the city is known to have one of the largest wastewater treatment plants in the world. Because the wastewater used by the large number of citizens flows into WWTP through the pipeline, the influent can contain variety of microbes as well as organic compounds. Since the microorganisms entering bioreactor in

¹School of Civil and Environmental Engineering, Yonsei University, Seoul, Republic of Korea. ²Department of Environmental Engineering, Yonsei University, Wonju, Republic of Korea. ³Present address: Department of Environmental Engineering, Korea Maritime and Ocean University, Busan, Republic of Korea. Young Kyung Kim and Keunje Yoo contributed equally. Correspondence and requests for materials should be addressed to T.K.L. (email: tklee@yonsei.ac.kr) or J.P. (email: parkj@yonsei.ac.kr)

WWTPs are mainly from the influent, microorganisms in the influent containing various substrates may influence community compositions or population diversity in the bioreactor¹². Meanwhile, some of newly developing Asian cities such as Ho Chi Minh or Hanoi started to manage relatively small-scale wastewater treatment facilities in small administrative districts or large buildings due to geographical limitations and construction/installation costs¹³. The biochemical composition of wastewater generated in such a small area is relatively homogeneous, and bacterial communities with a similar function are likely to be selected for biological reactors. The remarkable different capacities of the Seoul and Vietnamese WWTPs allow us to explore an interesting question, i.e., if bacterial composition in a bioreactor is influenced only by WWTP capacity or by other geographical factors.

The effects of capacity on bacterial communities in WWTPs were indirectly evaluated by comparison of 16S rRNA gene sequencing in previous studies. WWTPs in Hong Kong, which have a large WWTP capacity (>200,000 tonnes/day), showed significantly higher bacterial diversity compared to WWTPs in Singapore and the United States, which have relatively small in the capacity (8000–55,000 tonnes/day)⁶. Although Proteobacteria and Bacteroidetes at the phylum level dominant in all samples, the dominant genera showed geographical characteristics. Global studies on the diversity of bacterial communities in 269 WWTPs (23 countries on 6 continents) also observed that their geographical turnover is scale-dependent and driven by stochastic processes (dispersal and drift)¹⁴. WWTPs with larger capacity, the greater the diversity of microorganisms introduced, may increase the probability of stochastic processes occurring in community assemblies¹². Despite recent advances in our understanding of the effects of capacity on the bacterial diversity and structure of WWTPs, information on the bacterial interactions or community assembly remains elusive. This information is essential for identifying the key players in the biological treatment process and understanding the sludge assembly for stable wastewater treatment.

We collected AS samples from a total of eight WWTPs in Korea and Vietnam, which have similar operation parameters and treatment processes but different the capacities (flow rates). Deep sequencing of 16S rRNA gene and co-occurrence network analysis were performed to understand the effects of WWTP capacity on bacterial diversity and community assembly. We also evaluated the key players (hub members) to interpret their functional potentials in WWTPs with different capacities.

Materials and Methods

Sampling. Influent and AS from eight different WWTPs were used for this study. A total of 49 samples (8 influents and 41 AS) were collected from high-capacity WWTPs in Seoul, South Korea, between November 2014 and April 2015 (collectively referred to as 'KR_{WT}'). AS samples were obtained from anaerobic, anoxic and aerobic zones, respectively, in each WWTP and pooled into three or four AS samples to reduce potential sampling bias. The designed capacities and operating temperatures of these WWTPs ranged between 500,000 and 1,000,000 tonnes/day, and 10 °C and 16 °C, respectively. A total of 18 samples (7 influents and 11 AS) were collected from low-capacity WWTPs in Hanoi and Ho Chi Minh, Vietnam, between January and February 2015 (collectively referred to as 'VN_{WT}'). The designed capacities and operating temperatures of these WWTPs ranged between 1000 and 10,000 tons/day and 24 °C and 28 °C, respectively. The monthly average values for five water quality indicators [biological oxygen demand (BOD), chemical oxygen demand (COD), suspended solid (SS), total nitrogen (TN) and total phosphorus (TP)] of influent and effluent, were measured using environmental standard methods¹⁵. To take representative samples, sampling was performed at a minimum of 20 to 30 cm below the water surface. All samples were stored at –80 °C before extracting DNA.

DNA extraction and sequencing. Total genomic DNA from each AS was extracted using a FastDNA SPIN Kit for Soil (MP Biomedicals, Santa Ana, CA, USA) according to the manufacturer's protocol. DNA quality was checked using an ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA), and samples were stored at –20 °C.

A 16S amplicon polymerase chain reaction (PCR) forward and reverse primer targeting the V3 to V4 regions of the 16S rRNA gene¹⁶ was used to amplify the 16S rRNA gene from the extracted DNA. Amplification was conducted using a C1000TM thermal cycler (Bio-Rad, Hercules, CA, USA) with the following conditions: 94 °C for 1 min, 30 cycles at 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 2 min with a final extension at 72 °C for 5 min. PCR products were purified by gel electrophoresis bands isolated using a QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). The purified amplicons were sequenced by Macrogen Inc. (Seoul, Korea) using an Illumina Miseq sequencer. The data sets were deposited at the National Center for Biotechnology Information Sequence Read Archive database under accession number PRJNA533065.

Community analysis. Sequences were processed using MOTHUR software version 1.39.5¹⁷ according to the standard operating procedure (http://www.mothur.org/wiki/MiSeq_SOP)^{18,19} with minor modifications²⁰. Briefly, sequences longer than 466 bp were discarded, and those that remained were screened with start and end positions of 10,364 and 25,316, respectively, after being aligned to the SILVA database Release 132²¹. Taxonomy was classified using RDP 16S rRNA as a reference trainset 16²². The VSEARCH in MOTHUR was used to filter chimeras. Operational taxonomic units (OTUs) were clustered using the OptiClust algorithm with a 97% sequence similarity cutoff²³.

Network analysis. Co-occurrence networks were inferred based on a Spearman correlation matrix²⁴ and constructed using only significant correlation²⁵. The cutoff for correlation coefficients was determined to be 0.6 and the cutoff for p-values was 0.001²⁶. The cutoff was chosen by variance of interaction strength. All network construction steps were computed in R using code adapted from <https://github.com/ryanjw/co-occurrence>²⁷.

Statistical analysis. Statistical analysis was implemented using the R platform. The significance of the differences in bacterial diversity (richness and evenness) between WWTP types was evaluated by comparing Wilcoxon

WWTPs	KR1	KR2	KR3	KR4	VN1	VN2	VN3	VN4	p-value
Wastewater source	Domestic	Domestic	Domestic	Domestic	Domestic	Domestic	Domestic	Industry	—
Treatment process	MLE/A2O	MLE	MLE/A2O	MLE	MLE	MLE	A2O	SBR	—
avg SRT (days)	10	9	10	9	16	16	16	18	0.0247*
avg HRT (hours)	10	7	7	6	9	9	8	6	0.6704
avg Flow rate (m ³ /day)	850,000	1,500,000	1,700,000	900,000	1,000	1,300	1,000	18,000	0.0294*
avg MLSS (mg/L)	2,700	2,800	2,500	2,800	3,200	2,700	2,800	3,000	0.1379
pH	6.9	6.9	7.0	6.9	7.1	6.9	6.8	7.0	0.8770
COD _{rem} (%)	95.6	92.9	94.8	92.1	93.9	92.7	92.5	92.4	0.3123
BOD _{rem} (%)	92.6	91.7	94.3	92.9	92.5	93.4	91.8	93.6	0.9439
TN _{rem} (%)	91.8	96.8	90.4	93.9	97.7	92.7	95.5	96.3	0.2312
TP _{rem} (%)	97.1	97.8	96.5	92.3	99.8	99.0	92.9	97.9	0.4851

Table 1. Major operating parameters of WWTPs in this study. Asterisk indicates significant differences between KR_{WT} and VN_{WT}. MLE, Modified Ludzack-Ettinger; SBR, Sequencing Batch Reactor; A2O, Anaerobic/Anoxic/Oxic process.

rank-sum test p-values. Non-metric multidimensional scaling (NMDS) was performed using the ‘vegan’ package based on the relative abundance profiles at OTU levels. Water quality parameters were used for loading vectors contributing the ordination by permutation test (p-value < 0.05). Permutational multivariate analysis of variance (PERMANOVA) was performed using the ‘vegan’ package, and the permuted p-value was obtained by 10,000 times of permutations.

Results and Discussion

Characteristics of WWTPs in Korea and Vietnam. Table 1 summarises the operating parameters and removal efficiencies for each of the eight WWTPs in Korea and Vietnam. Most of biological treatments of WWTPs in this study have been operated by anaerobic/anoxic/oxic (A2O) or modified Ludzack-Ettinger (MLE) processes for treating domestic wastewater, except for one WWTP that used sequencing batch reactor (SBR) for industry wastewater. AS VN_{WT} in this study was constructed by Korean construction companies, similar processes as KR_{WT} were applied. Regardless of country and region, WWTPs had similar overall operating parameters and removal efficiencies. Flow rates and SRT were significantly different between countries (p-value < 0.05) whereas other parameters (HRT, mixed liquor SS [MLSS] and pH) were comparable to each other. The differences in WWTPs in the two countries can be explained by historical and social reasons. When Seoul was developed in earnest in the 1970s, residential plans and WWTP construction plans were established spontaneously. The pipe construction and WWTP expansion plans could be considered at the same time; thus, the size of KR_{WT} was successfully expanded as to handle approximately 1,000,000 m³/day due to the explosive increase in the population of Seoul. However, Vietnam had many buildings already built in the city, so it was difficult to build large WWTPs due to geographical limitations and construction/installation costs¹⁵. VN_{WT} has recently to build on a small scale because WWTPs have been built in a large new building or a small town. SRT is closely related to the flow rate because it is affected by the amount of sludge generated by the capacity of the WWTPs and the required MLSS concentrations. Taken together, the flow rates of the two regional WWTPs are considered to have the largest difference in operating parameters.

The water quality parameters of influents and effluents were monitored to compare treatment performances for KR_{WT} and VN_{WT} samples (Supplementary Fig. S1). The influent composition exhibited no significant difference between the two countries, except that KR_{WT} had a wider distribution in BOD and COD and a narrower distribution in SS and TN. The wider distribution of BOD and COD suggested that there are a various types of substrates entering wastewater or a conversion of organic matter during travel of wastewater from the source to WWTP²⁸. Such a trend was absent in VN_{WT} samples because the contaminants produced in the same facilities or small town were always similar and similar types of substrates flowed into the influent. These results indicated that WWTPs with large capacity are likely to have various COD and BOD. The high fluctuation in influent TN and SS could be the unique characteristic of small WWTPs, as individuals have greater contribution to the composition of wastewater. Thus, changes in water usage patterns of a small portion of inhabitants can result in large change in influent characteristics. For instance, about 30% to 40% of household wastewater is from bathrooms²⁹ and because small WWTPs serve fewer households compared to the large WWTPs, minor changes in the pattern of individuals occupying bathrooms at any given hour can have a large effect on the influent quality of small WWTPs. The high COD of VN_{WT} samples suggests different design criteria of WWTPs and/or AS less climatized for degrading recalcitrant organic substances which requires a series of complex microbial actions³⁰.

Bacterial diversity and composition. Sequencing obtained 27,581 ± 10,031 reads and 28,915 ± 10,535 reads of quality reads from influent and AS of KR_{WT} and VN_{WT}, respectively. Good’s coverages were all above 95%, indicating that the number of sequences was enough for sampling the communities. The richness (observed OTUs) of KR_{WT} influent exhibited less variable compared to that of VN_{WT} influent (Fig. 1A). Such a large spread could be due to the fact that minor change in water usage can have a large effect on influent characteristics for small WWTPs, as discussed in Section 3.1. The richness increased in AS for both WWTP types with comparable variances. Nevertheless, a larger increase was observed for KR_{WT} AS (p-value < 0.001; average richness = 546), whereas the increase for VN_{WT} AS was less (p-value < 0.05; average richness = 373). The evenness of KR_{WT} and

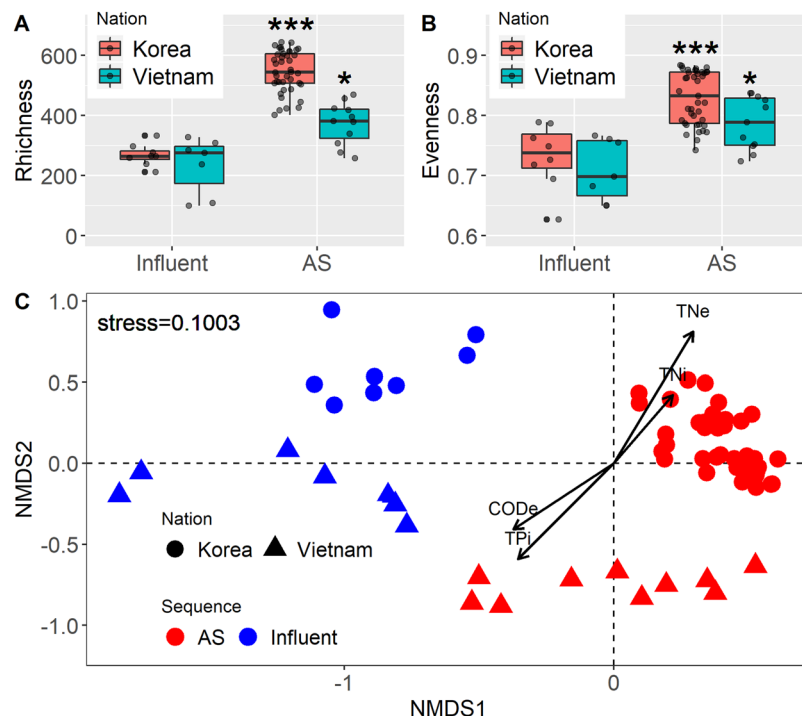


Figure 1. Comparison of microbial diversity and community structure between WWTPs in Korean and Vietnam. **(A)** Richness (number of observed OTUs), **(B)** evenness, **(C)** NMDS plot showing microbial communities present in this study ($k = 2$; stress = 0.1003). Water quality parameters indicated by arrows are significant parameters (p -value < 0.05) contributing to each ordination. CODE, COD in effluent; TNe, TN in effluent; TNi, TN in influent; TPi, TP in influent.

VN_{WT} exhibited similar trends to the richness (Fig. 1B). It is widely recognised that parameters such as pH, total organic carbon, essential nutrients, and dissolved oxygen (DO) have a significant impact on bacterial communities in various ecosystems^{31,32}. Thus, bacterial diversity and evenness may reflect not only the WWTP design but also influent characteristics, which are strongly affected by the cultural, social and environmental differences in each region³². Gao *et al.*³³ reported that COD and pH are the primary variables of influent that can affect the bacterial community of AS. Although pH data were unavailable in this study, COD was one of the parameters that exhibited a large variation in KR_{WT} influent. The large variability of COD can be linked to the influx of various types of substrates in the influent, which might be manifested by increasing the diversity of microorganisms depending on the substrate specificity.

The differential distribution of bacterial populations between KR_{WT} and VN_{WT} on an NMDS plot was statistically confirmed using PERMANOVA ($R = 0.786$, p -value < 0.05 ; Fig. 1C). In general, the y-axis separated the sequences based on their origin, with KR_{WT} on the positive side and VN_{WT} on the negative side of the plot, whereas the x-axis separated sequences based on the wastewater treatment phase, with more negative values for the influent than the AS. The water quality parameters including (TNe: TN in effluent, TNi: TN in influent, CODE: COD in effluent, TPi: TP in influent) were selected as significant loading vectors (p -value < 0.05) contributing the ordination in the NMDS plots. In previous studies, several WWTPs processing influents with differing characteristics and using different operating parameters differed significantly in their bacterial community structures^{6,32,34,35}. However, in this study, the bacterial communities found at the same wastewater treatment phase were more related than those from the same location. This is reasonable because although KR_{WT} and VN_{WT} influents have different characteristics, and undergo different biochemical reactions in the sewer network, this does not change the fact that they are generated from human activities. Furthermore, two of VN_{WT} were from WWTPs designed for in-complex wastewater treatment. Thus, as the influent was always composed of similar types or components of pollutants due to the limited source of influent in the in-complex WWTP, the composition of bacterial communities might be specifically formed according to each WWTP in Vietnam. Likewise, the similar bacterial communities of AS also contributed to the clustering of the samples from different locations. Although the designed capacities differed at each WWTP, their main purpose is the same – that is, the removal of organic matter and excess nutrients. Thus, although the designs for WWTPs of KR_{WT} and VN_{WT} samples may be different, both types were built based on well-established biological methods that rely strongly on the development of specific bacterial communities for their performance. Notwithstanding the numerous factors that could contribute to the differences in AS bacterial communities, such as treatment capacity, operation parameters, and local environmental conditions, the shift was insufficient to break the clustering trend.

In this study, Proteobacteria was the most abundant phylum in all samples of both KR_{WT} and VN_{WT}, accounting for 40% to 60% of the total effective bacterial sequences (Fig. 2). This is consistent with other results of bacterial communities in influent³⁶ and AS^{6,37}. The other substantial phyla identified in this study included Bacteroidetes

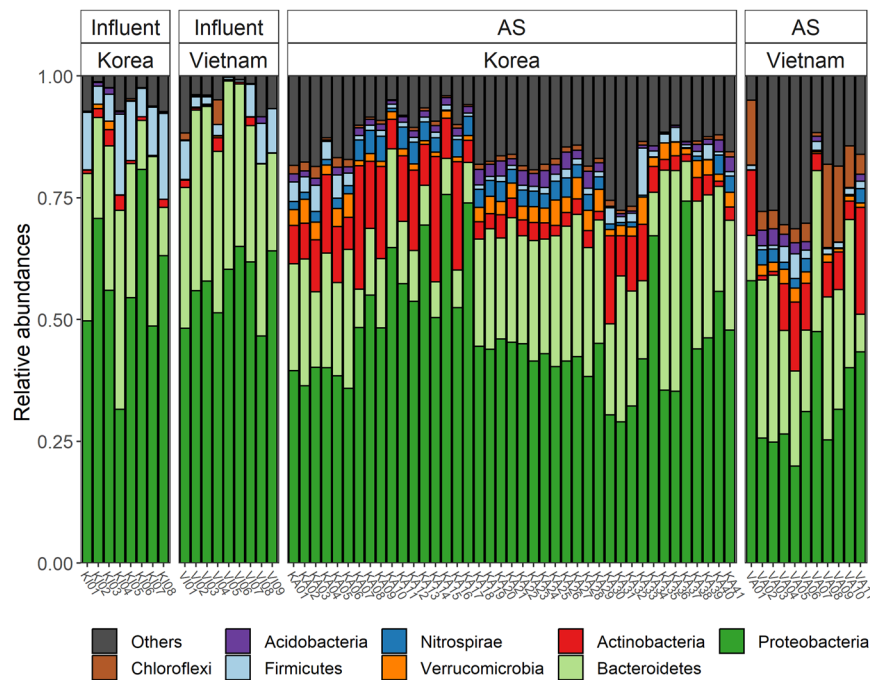


Figure 2. Phylum level microbial community composition in influents and WWTPs.

(20–30%), Firmicutes (5–10%), and Actinobacteria (3–8%). Proteobacteria are believed to be involved in the removal of organic pollutants, such as nitrogen, phosphorus, and aromatic compounds^{5,37}. Wagne and Loy⁵ surveyed eight different WWTPs in which Beta-, Alpha-, and Gammaproteobacteria, as well as Bacteroidetes, were frequently found, in that order, although the relative ratios were somewhat variable. Previous studies on AS using microarrays⁸ or cloning¹ similarly found that the same four groups identified in this study were dominant (around 80–85%) in bacterial communities, followed by a few other major (average abundance >1%) phyla, including Chloroflexi, Nitrospirae, Acidobacteria, Candidatus Saccharibacteria, and Verrucomicrobia. A few phyla in this study, such as Bacteroidetes, Firmicutes, Actinobacteria, Nitrospirae, Candidatus Saccharibacteri, and Verrucomicrobia, differed significantly between Korean and Vietnamese AS samples (p -value < 0.05).

These results suggest that bacterial communities develop differently depending on the capacity of WWTP, which is consistent with previous studies^{3,4}. Wastewater engineers have tried to develop treatment processes by creating specific conditions for harnessing particular microorganisms^{6,38}. Specific environmental conditions, including abiotic factors (e.g. temperature, DO, pH, organic matter, HRT and SRT) and interspecies interactions (e.g. competition, symbiosis and predation), may influence on the selection of specific bacterial populations from the influent wastewater community that are better suited to the bioreactor. Then, specific niches to the capacity of WWTP are created with low dispersal rate, and significant different communities could be assembled in the bioreactors. Previous studies suggested that the geographical distribution of WWTPs might be an important and influential factor in the variation of WWTP bacterial communities^{32,34,35}.

Network analysis. Diversified functional traits are closely associated with ecological niche differentiation, and network topology allows us to focus on the patterns of bacterial relationships that develop within a given ecosystem³⁹. Network analysis was incorporated to explore the assembly of bacterial communities in the high-capacity KR_{WT} and low-capacity VN_{WT} (Fig. 3). KR_{WT} samples had more nodes (98) and edges (287) as well as higher network density ($D = 0.0604$) and transitivity ($T = 0.4891$) compared to VN_{WT} samples (nodes = 82, edges = 146, $D = 0.0440$, and $T = 0.4482$). These results indicated that microorganisms from high-capacity WWTPs are more interrelated than those from low-capacity WWTPs, which correspond to the tighter clustering of KR_{WT} in NMDS (Fig. 1). These results indicated cooperation among diverse microorganisms in AS for wastewater treatment. Furthermore, this interaction was broader and richer in large-capacity WWTPs, which could be foreseen from the significantly higher diversity and evenness of KR_{WT} compared to that of VN_{WT} (Fig. 1). These results suggested that higher wastewater treatment capacity promotes bacterial diversity and interaction in AS. High-capacity WWTPs serving larger populations receive more diverse organic matter and nutrients from more sources than the low-capacity WWTPs. This facilitates not only broader metabolic reactions that utilise diverse substrates found in wastewater but also the expansion of ecological niches responsible for bioconversion⁴⁰. This competition or mutualistic interaction might contribute to stimulating taxonomic and functional diversity⁴¹. Frequent physicochemical and nutritional variation in a single ecosystem strongly influences bacterial community structure, but co-occurrence network patterns are quite persistent and interactive because of ecological rearrangements, in such a way that an important fraction of OTUs come to fill the interactive roles occupied by other OTUs in the network when facing strong environmental changes⁴².

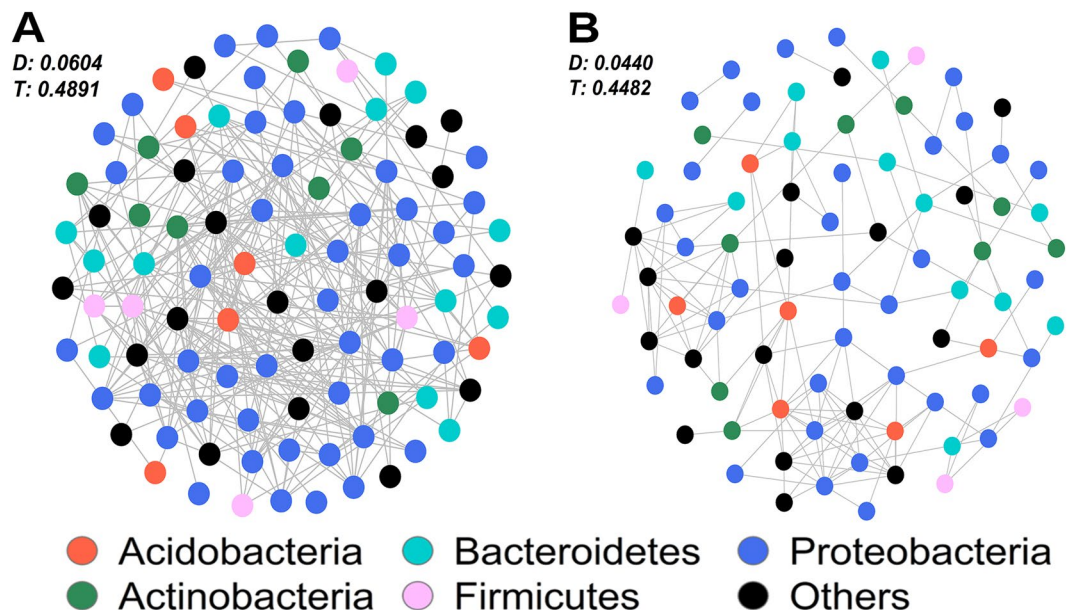


Figure 3. Network of co-occurring microbial OTUs of AS based on correlation analysis for (A) Korea (nodes = 98, edges = 287) and (B) Vietnam (nodes = 82, edge = 146). A connection stands for a strong (Spearman's $\rho > 0.6$) and significant (p -value < 0.001) correlation. Nodes are colored according to phylum. D, Density; T, Transitivity.

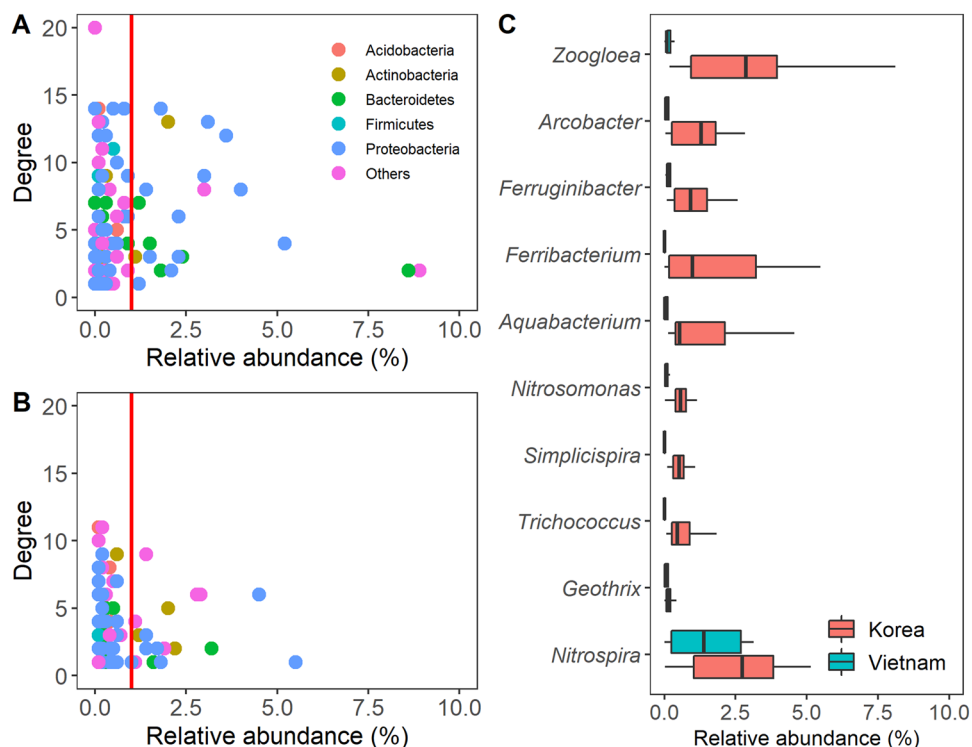


Figure 4. Relationship between the number of degree and relative abundance of each node (OTU) in AS for Korea (A) and Vietnam (B). The dominant phyla members are highlighted. (C) Relative abundances of the members with high degree in this study.

Hub members were defined as microbes with a high degree (number of connections) of co-occurrence (top 5%) and those are key players in sustaining the overall community⁴³. More hubs were observed in KR_{WT} (10 hubs) than VN_{WT} (6 hubs), but most of the detected hubs were of relatively low abundance ($< 1\%$ of total), particularly those in VN_{WT}, which were close to zero (Fig. 4A,B). It is worth noting that KR_{WT} had a significantly

higher number of hub-like members (>3 degrees) with abundances greater than 1%, but VN_{WT} members with abundances of more than 1% were very rare. Meanwhile, hubs in VN_{WT} were assigned to others (non-dominant phyla in this study) or Acidobacteria, whereas those in KR_{WT} were mainly affiliated to diverse phyla, including Acidobacteria, Actinobacteria, Firmicutes, and Proteobacteria. These results indicate that diverse microorganisms in KR_{WT} tended to be metabolically active and established heterogeneous interactions with other microorganisms that aided wastewater treatment. Bacterial diversity has been shown to be a major factor associated with network complexity, suggesting induced niche differentiation, functional redundancy, and process stability⁴⁴. These may result in the maintenance of performance recovery, operational stability, and bacterial flexibility in KR_{WT} being more resilient to common disturbances in influent characteristics¹¹.

Most members with high degree had a significantly higher relative abundance in Korea compared to Vietnam (p-value < 0.05), except for *Nitrospira* and *Geothrix* (Fig. 4C). Hub members have been shown to significantly correlate with the function and performance of wastewater treatment³⁷; thus, the members could serve as biological indicators for each region. Hub members were commonly found in other WWTPs, as described in previous studies^{34,35}. For instance, *Zoogloea* and *Trichococcus* are the main agents for the flocculation of AS because they form extracellular gelatinous matrices^{34,45}, whereas *Ferruginibacter* and *Arcobacter* are capable of hydrolysing some organic matter⁴⁶. *Nitrosomonas* plays a core role in nitrification in WWTPs⁴⁷. Interactions between microorganisms and members with high degree may perform key biogeochemical processes (e.g., organic removal, nutrient mineralisation, nitrification and phosphorus removal) and stabilise sludge flocs. Such information can provide a basis for targeted manipulation of AS community for the stable and efficient operation of the WWTPs. If AS can be manipulated, the efficiency of wastewater treatment can be increased with reduced treatment time, and the function of sludge can recover quickly when the sludge is damaged by external disturbances.

Conclusion

This study provides evidence of the impact of capacity on bacterial diversity and community assembly in WWTPs. Although the capacity itself is not directly related to bacterial communities and relationships, changes in COD, SRT and other conditions depending on capacity, could indirectly influence bacterial diversity and its assembly. We have shown that high-capacity WWTPs contain relatively high bacterial diversity and that their bacterial communities are more interrelated than those of low-capacity WWTPs, which results to a tighter clustering of KR_{WT} in NMDS. These differences seem to be induced by the probability of stochastic processes in the AS assembly in different capacities of WWTPs. Co-occurrence network analysis provides information on key players in AS, which may be the basis for manipulating the AS stability in biological treatment processes. The findings support the hypothesis that capacity is a key factor influencing AS bacterial community and its assembly, providing a fundamental bioreactor engineering implication to optimize AS depending on capacity of WWTPs.

References

- Snaird, J., Amann, R., Huber, I., Ludwig, W. & Schleifer, K. H. Phylogenetic analysis and *in situ* identification of bacteria in activated sludge. *Appl. Environ. Microbiol.* **63**, 2884–2896 (1997).
- Wells, G. F. *et al.* Ammonia-oxidizing communities in a highly aerated full-scale activated sludge bioreactor: betaproteobacterial dynamics and low relative abundance of Crenarchaea. *Environ. Microbiol.* **11**, 2310–2328 (2009).
- Ibarbalz, F. M., Figuerola, E. L. M. & Erijman, L. Industrial activated sludge exhibit unique bacterial community composition at high taxonomic ranks. *Water Res.* **47**, 3854–3864 (2013).
- Saunders, A. M., Albertsen, M., Vollertsen, J. & Nielsen, P. H. The activated sludge ecosystem contains a core community of abundant organisms. *ISME J.* **10**, 11 (2015).
- Wagner, M. & Loy, A. Bacterial community composition and function in sewage treatment systems. *Curr. Opin. Biotech.* **13**, 218–227 (2002).
- Zhang, T., Shao, M. F. & Ye, L. 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *ISME J* **6**, 1137–1147 (2012).
- Barr, T. A., Taylor, J. M. & Duff, S. J. B. Effect of HRT, SRT and temperature on the performance of activated sludge reactors treating bleached kraft mill effluent. *Water Res.* **30**, 799–810 (1996).
- Xia, S. *et al.* Bacterial community structure in geographically distributed biological wastewater treatment reactors. *Environ. Sci. Technol.* **44**, 7391–7396 (2010).
- Park, H.-D. & Noguera, D. R. Evaluating the effect of dissolved oxygen on ammonia-oxidizing bacterial communities in activated sludge. *Water Res.* **38**, 3275–3286 (2004).
- Nowak, O., Kuehn, V. & Zessner, M. Sludge management of small water and wastewater treatment plants. *Water Sci. Technol.* **48**, 33–41 (2003).
- De Vrieze, J. *et al.* Microbial community redundancy in anaerobic digestion drives process recovery after salinity exposure. *Water Res.* **111**, 109–117 (2017).
- Zhang, B., Xu, X. & Zhu, L. Structure and function of the microbial consortia of activated sludge in typical municipal wastewater treatment plants in winter. *Sci. Rep.* **7**, 17930, <https://doi.org/10.1038/s41598-017-17743-x> (2017).
- World Bank. Vietnam urban wastewater review: Executive summary, <http://documents.worldbank.org/curated/en/385401468262139190/Executive-summary> (2013).
- Wu, L. *et al.* Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nat. Microbiol.* **4**, 1183–1195 (2019).
- Ministry of Environment. Standard methods for the examination of water pollution, <http://www.me.go.kr> (2017).
- Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1–e1, <https://doi.org/10.1093/nar/gks808> (2013).
- Schloss, P. D. *et al.* Introducing Mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537 (2009).
- Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112 (2013).
- Lopez-Garcia, A. *et al.* Comparison of Mothur and QIIME for the analysis of rumen microbiota composition based on 16S rRNA amplicon sequences. *Front. Microbiol.* **9**, 3010, <https://doi.org/10.3389/fmicb.2018.03010> (2018).
- Yoo, K., Yoo, H., Lee, J. M., Shukla, S. K. & Park, J. Classification and regression tree approach for prediction of potential hazards of urban airborne bacteria during Asian dust events. *Sci. Rep.* **8**, 11823, <https://doi.org/10.1038/s41598-018-29796-7> (2018).

21. Quast, C. *et al.* 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> (2012).
22. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
23. Westcott, S. L. & Schloss, P. D. OptiClust, an improved method for assigning amplicon-based sequence data to operational taxonomic units. *mSphere* **2**, e00073–00017, <https://doi.org/10.1128/mSphereDirect.00073-17> (2017).
24. Langfelder, P. & Horvath, S. Fast R functions for robust correlations and hierarchical clustering. *J. Stat. Softw.* **46**, i11, <https://doi.org/10.18637/jss.v046.i11> (2012).
25. Barberan, A. *et al.* Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. *Ecol. Lett.* **17**, 794–802 (2014).
26. Luo, X. & Bhattacharya, C. B. Corporate social responsibility, customer satisfaction, and market value. *J. Mark.* **70**, 1–18 (2006).
27. De Vries, F. T. *et al.* Soil bacterial networks are less stable under drought than fungal networks. *Nat. Commun.* **9**, 3033–3033, <https://doi.org/10.1038/s41467-018-05516-7> (2018).
28. McCall, A.-K. *et al.* Critical review on the stability of illicit drugs in sewers and wastewater samples. *Water Res.* **88**, 933–947 (2016).
29. Butler, D., Friedler, E. & Gatt, K. Characterising the quantity and quality of domestic wastewater inflows. *Wat. Sci. Tech.* **31**, 13–24 (1995).
30. Choi, Y.-Y. *et al.* Characteristics and biodegradability of wastewater organic matter in municipal wastewater treatment plants collecting domestic wastewater and industrial discharge. *Water* **9**, 409 (2017).
31. Amann, R., Lemmer, H. & Wagner, M. Monitoring the community structure of wastewater treatment plants: a comparison of old and new techniques. *FEMS Microbiol. Ecol.* **25**, 205–215 (1998).
32. Wang, X., Xia, Y., Wen, X., Yang, Y. & Zhou, J. Microbial community functional structures in wastewater treatment plants as characterized by geochip. *PLoS One* **9**, e93422, <https://doi.org/10.1371/journal.pone.0093422> (2014).
33. Gao, P. *et al.* Correlating microbial community compositions with environmental factors in activated sludge from four full-scale municipal wastewater treatment plants in Shanghai, China. *Appl. Microbiol. Biotechnol.* **100**, 4663–4673 (2016).
34. Wang, X., Hu, M., Xia, Y., Wen, X. & Ding, K. Pyrosequencing analysis of bacterial diversity in 14 wastewater treatment systems in China. *Appl. Environ. Microbiol.* **78**, 7042 (2012).
35. Zhao, D. *et al.* Pyrosequencing analysis of bacterial community and assembly in activated sludge samples from different geographic regions in China. *Appl. Microbiol. Biotechnol.* **98**, 9119–9128 (2014).
36. McClellan, J. & King, M.-C. Genetic heterogeneity in human disease. *Cell* **141**, 210–217 (2010).
37. Yang, C. *et al.* Phylogenetic diversity and metabolic potential of activated Sludge microbial communities in full-scale wastewater treatment plants. *Environ. Sci. Technol.* **45**, 7408–7415 (2011).
38. Ibarbalz, F. M., Orellana, E., Figuerola, E. L. M. & Erijman, L. Shotgun metagenomic profiles have a high capacity to discriminate samples of activated sludge according to wastewater type. *Appl. Environ. Microbiol.* **82**, 5186 (2016).
39. Cumming, G. S., Bodin, Ö., Ernstson, H. & Elmqvist, T. Network analysis in conservation biogeography: challenges and opportunities. *Divers. Distrib.* **16**, 414–425 (2010).
40. Lee, S.-H., Kang, H.-J. & Park, H.-D. Influence of influent wastewater communities on temporal variation of activated sludge communities. *Water Res.* **73**, 132–144 (2015).
41. Meyer, J. R. & Kassen, R. The effects of competition and predation on diversification in a model adaptive radiation. *Nature* **446**, 432–435 (2007).
42. Mandakovic, D. *et al.* Structure and co-occurrence patterns in microbial communities under acute environmental stress reveal ecological factors fostering resilience. *Sci. Rep.* **8**, 5875, <https://doi.org/10.1038/s41598-018-23931-0> (2018).
43. Berry, D. & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.* **5**, <https://doi.org/10.3389/fmicb.2014.00219> (2014).
44. Wu, L. *et al.* Long-term successional dynamics of microbial association networks in anaerobic digestion processes. *Water Res.* **104**, 1–10 (2016).
45. Xie, C. H. & Yokota, A. *Zoogloea oryzae* sp. nov., a nitrogen-fixing bacterium isolated from rice paddy soil, and reclassification of the strain ATCC 19623 as *Crabtreeella saccharophila* gen. nov., sp. nov. *Int. J. Syst. Evol. Microbiol.* **56**, 619–624 (2006).
46. Lim, J. H., Baek, S. H. & Lee, S. T. *Ferruginibacter alkalilentus* gen. nov., sp. nov. and *Ferruginibacter lapsinensis* sp. nov., novel members of the family ‘Chitinophagaceae’ in the phylum Bacteroidetes, isolated from freshwater sediment. *Int. J. Syst. Evol. Microbiol.* **59**, 2394–2399 (2009).
47. Purkhold, U. *et al.* Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and *amoA* sequence analysis: implications for molecular diversity surveys. *Appl. Environ. Microbiol.* **66**, 5368–5382 (2000).

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (NRF-2017R1D1A3B03029787) and the Korea Institute of Energy Technology Evaluation and Planning and the Ministry of Trade, Industry & Energy of the Republic of Korea (No. 20184030202240).

Author Contributions

Y.K.K. and K.Y. performed data analysis and wrote the manuscript. M.L., M.S.K. and T.K.L. performed laboratory work and processed sequencing data. T.K.L. and J.P. designed the study. T.K.L. and J.P. were involved in the organisation and management of the overall project. I.H., B.R.K. and T.K.L. were involved in the peer editing of the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-50952-0>.

Competing Interests: The authors declare no competing interests.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019