Evaluation of influent microbial immigration to activated sludge is affected by different-sized community segregation

Bing Guo, Zhiya Sheng and Yang Liu

Activated sludge (AS) microbial communities were analyzed for seasonal variation, a disturbance-recovery event, and separated small aggregates (SAG) to study the influent immigration effect using both neutral immigration model and mass-balance model with operational parameters. SAG differed with AS, and higher immigration impact on SAG was confirmed by both models. Adding the SAG community segregation in the latter model to evaluate the contribution of influent immigration to community disturbance-recovery showed increased impact of immigration.

npj Clean Water (2021) 4:20; https://doi.org/10.1038/s41545-021-00112-7

INTRODUCTION

The activated sludge (AS) wastewater treatment plant provides a well-defined model system to study microbial ecology. The AS community is shaped by deterministic (Niche) factors and stochastic (Neutral) factors (e.g., neutral immigration from wastewater influent)\(^1\). Two types of numerical models have been used to quantify the influent immigration effect. One is the neutral community (immigration) model (NIM)\(^2\), employing the abundance and frequency of operational taxonomic units (OTUs) in metacommunities to calculate the community-level migration probability\(^3\)\(^4\)\(^5\). The other is mass-balance model (MBM) that defines the immigration as the proportion of reactor biomass derived from influent biomass \(\text{influent biomass} \times \frac{\text{influent biomass}}{\text{influent biomass} + \text{local biomass}}\), using abundances of specific OTUs to calculate the immigration rates and net growth rates\(^6\). The NIM immigration describes the probability that a dead individual in the reactor is replaced by an immigrant from the influent, differentiated from the reproduction of local community. Intrinsically, the MBM bases on the fractionation of influent biomass and local growth, which also aims to differentiate the influent immigration and local community. The MBM has been increasingly used because of the quantitative measurements of the OTU growth activity\(^7\). Although steady-state metacommunities demonstrated that operational parameters (e.g., solid retention time, SRT) are important for the evaluation of immigration\(^8\)\(^9\), the operation parameters were not directly employed in the original MBM. Recently, Frigon and Wells\(^10\) and Guo\(^11\) developed the mass-flow immigration model, which is a MBM incorporating operational parameters (MBM-OP), to calculate the immigration rate of specific OTUs. By using operational parameters, this model is comprehensible for both wastewater engineers and microbiologists.

In wastewater treatment bioreactors, community variation has been shown in relation to the size and status (settled and unsettled) of the aggregates\(^12\)\(^13\). In conventional AS, the relative importance of influent immigration on different-sized aggregate communities (i.e., easily settleable large-size flocs and low-settleability small-size particles) has not been explored. Besides steady-state modeling, the bioreactors often experience disturbances\(^14\), such as sudden high flow rates which lead to a biomass washout. The non-steady-state community serves as a good example to study the community recovery and to compare the contributions of local community growth and influent immigration.

Our study aimed at monitoring the long-term variation of AS communities in different seasons and comparing the influent immigration effect on different-sized AS communities using NIM and MBM-OP models. In addition, the short-term non-steady-state (under operational high-flow disturbance) AS community’s recovery attributed to local community growth and influent immigration was evaluated.

AS community variation

The AS communities (between 2013 and 2019) showed that abundant microorganisms were relatively consistent during the long-term observation, and consisted of a core group at the order level (Fig. 1b) and at genus level (Supplementary Fig. S1). Seasonal changes and disturbance did not alter the core group\(^15\). The variation between small aggregates (SAG) and AS was clearly shown (Fig. 1b), being consistent in different seasons. The most abundant members in SAG were different from AS communities, but similar with the influent community’s most abundant members, indicating that influent microbiome is an important origin for the SAG. The community richness (number of observed OTUs) was slightly higher in SAG than in AS (Supplementary Fig. S2), inferring that the SAG community was a union of AS and influent communities. The shared numbers and abundances of OTUs are indicators of immigration effect\(^7\)\(^9\)\(^15\). A large number of OTUs were shared among the influent, AS, and SAG (Fig. 1c). The shared OTUs between influent and AS took up 89.1 ± 0.1% of the influent community total number of reads and 54.0 ± 4.3% of the AS community total number of reads, which increased between influent and SAG (97.0 ± 0.1% and 73.8 ± 3.0%, respectively), indicating higher immigration effect for influent–SAG than for influent–AS. The Bray-Curtis distances also showed higher similarity between influent and SAG than between influent and AS (Fig. 1d). Permutational multivariate analysis of variance (Bray-Curtis distance, permutation = 999) was performed to test the two-way factor effect of sampling time and aggregate size. Our results showed significant effect of aggregated size (\(p = 0.001\)), but not sampling time (\(p = 0.911\)) or their interaction (\(p = 0.866\)).

This phenomenon of community segregation was observed in different forms of biomass in biofilm and granular sludge.
activity (Supplementary Fig. S3)14,18,19. has higher average towards higher immigration rate compared to in growth) (Fig. 2a). A conceptual model MBM-OP-SAG was constructed (Fig. 2c). In both models, OTUs were classified into three groups based on their relative abundances in the influent and in the AS: (i) not in influent but in reactor community, (ii) growth group (calculated net growth rate $\mu_{\text{MBM-OP}} > 0$, Eq. 16), (iii) complete immigration group (calculated net growth rate $\mu_{\text{MBM-OP}} \leq 0$, Eq. 16). The relative abundances (Fig. 2b, d) showed that the immigration group was using the MBM-OP-SAG model more (Fig. 2d) than using the MBM-OP-AS model (Fig. 2b). For the disturbed and recovered AS communities, the MBM-OP-AS model showed similar abundances of immigration group (Fig. 2d). The MBM-OP-SAG model estimated higher contribution of immigration in recovered than in disturbed community (Fig. 2d). The two models differed in certain taxon’s growth rates and immigration rates, resulting in a list of genera classified into growth group by MBM-OP-AS model while into immigration group by MBM-OP-SAG model (Supplementary Table S1). Microorganisms such as Methanobacterium, Methanobrevibacter, Clostridiales, Synergistales, and Bifidobacterium are known anaerobic microorganisms and likely to be carried by wastewater from sewer20 and not to grow in AS. The MBM-OP-SAG model was successful to identify these genera as influential immigration microorganisms.

Our results demonstrated AS community variation in different-sized aggregates. NIM and MBM-OP models were compared, both resulted in higher immigration impact for the small aggregate (SAG) community than for the AS community. The different-sized community variation led to an approach (MBM-OP-SAG model) of modeling influent and segregated communities for immigration effects, which was applied to investigate the community recovery. The MBM-OP-SAG model showed higher immigration contribution to recovery than the MBM-OP-AS model. This approach could be applied to other bioreactor systems to increase modeling accuracy of microbial growth and immigration, compared to the conventional NIM and MBM-OP models.

Immigration contribution to community recovery

The AS community was disturbed by a high-flow event which led to a drop in biomass density, shifting away from the steady-state community and recovered after 3 days (Fig. 1d). The recovery could be attributed to different mechanisms, including growth of the indigenous microorganisms5,15 and influent immigration1,2,9.

Using the MBM-OP immigration model, the influent community can be partitioned to growth and neutral immigration (non-growth) (Fig. 2a). A conceptual model MBM-OP-SAG was constructed (Fig. 2c). In both models, OTUs were classified into
DNA extraction, 16S rRNA gene amplicon sequencing

DNA was extracted from influent and bioreactor samples using DNeasy PowerSoil Kit (Qiagen Inc., Toronto, Canada) according to the manufacturer’s protocol. The long-term monitoring of the community was conducted using Illumina MiSeq PE250 platform. Sequence data were deposited to the National Center for Biotechnology Information (NCBI) GenBank (Bio Project Accession number PRJNA580466).

The raw sequences (2 x 250 nt) were paired, quality-filtered using DADA2 in QIIME2 pipelines. The OTUs were defined using the default 100% similarity. The taxonomic identification was assigned using GreenGenes (version gg_13.8) reference database at 99% similarity. Microbial community diversities (alpha-diversity, beta-diversity) were analyzed using R “vegan” package. The weighted rm copy number was estimated by the normalized OTU table using PICRUSt with metagenome references in Integrated Microbial Genomes (IMG) system.

Weighted rm copy number = \sum_{i=1}^{n_{16S_norm}} N_{r{i}} 

Neutral immigration model

The probability of influence of immigration was defined by Sloan et al., and calculated using the neutral immigration model. The model described that in a microbial community with Ni individuals, an individual die or leave the system and is immediately replaced by an immigrant with probability of m, or by reproduction of a local member of the community with probability of 1 – m.

For the i-th species (initial number Ni), the probability (Pr) of increase by one, stay the same, or decrease by one is expressed, respectively, by the following three equations:

\[ Pr\left(N_i + \frac{1}{N_i}\right) = \frac{N_i - N_i}{N_i} \left[m \cdot \left(1 + \alpha_i\right) \left(1 - m\right) \left(\frac{N_i}{N_i - 1}\right)\right] \]

\[ Pr\left(N_i / N_i\right) = \frac{N_i}{N_i} \left[m \cdot \left(1 - \alpha_i\right) \left(1 - m\right) \left(\frac{N_i}{N_i - 1}\right)\right] \]

\[ Pr\left(N_i - \frac{1}{N_i}\right) = \frac{N_i}{N_i} \left[\left(1 - \alpha_i\right) \left(1 - m\right) \left(\frac{N_i}{N_i - 1}\right)\right] \]

where \( \alpha_i \) is the advantage of the i-th species over the other species.
where $c$ is a constant so that $\frac{1}{d} \int_{d} |x| \, dx = 1$.

The probability of the $i$th species observed in any local community is:

$$P_r(species \ i \ is \ present \ with \ a \ relative \ abundance > d) = \int_{d} \phi(\psi) \, dx$$

where $d$ is the detection threshold.

This equation can be used to calibrate community data (frequency and mean relative abundance plot) and solve for $m$.

$AS$ and $SAG$ communities included six replicates each. The immigration rate was estimated for influent and $AS$ communities and influent and $SAG$ communities, using the $R$ codes from Burns et al$^2$, with 100 permutations ($R$ code in Supplementary Information).

### Mass-balance model with operational parameters

The MBM with operational parameters (mass-flow immigration model) calculation for net growth rates was described by Guo$^1$.

The mixed liquor heterotrophic biomass is contributed from the influent biomass and the microbial growth from the consumption of substrate resources. The immigration is defined as the fraction of influent biomass contribution:

$$m = \frac{\text{influent biomass}}{\text{influent biomass} + \text{local growth}} = \frac{f_{\text{HO,Capt}} + f_{\text{HO,inf}}}{f_{\text{HO,Capt}} + f_{\text{HO,inf}} + f_{\text{ML,inf}} + f_{\text{consumed}}}$$

where $m$ is the mass-flow immigration rate; $f_{\text{HO,inf}}$ and $f_{\text{HO,ML,inf}}$ are the active heterotrophic biomass in influent and mixed liquor; $f_{\text{HO,Capt}}$ is the fraction of influent biomass captured by the mixed liquor which is default $1$; $f_{\text{ML,inf}}$ is the heterotrophic growth yield; $f_{\text{consumed}}$ is the consumed substrates.

For continuous stirred tank reactors, the heterotrophic biomass based on mass-balance calculation$^{20}$ is:

$$X_{\text{HO,ML}} = \frac{\theta_x}{\delta} \left( \frac{Y_{\text{HO}}}{1 + \theta_{\text{HO}} + \theta_x} \right) + \frac{\theta_y}{\delta} \left( \frac{S_{\text{consumed}}}{1 + \theta_{\text{ML,inf}} + \theta_y} \right)$$

Rearrangement gives:

$$Y_{\text{HO}} + S_{\text{consumed}} = X_{\text{HO,ML}} (\theta_x/\delta)/(1 + \theta_{\text{HO}} + \theta_x)$$

Under the assumption that $f_{\text{HO,Capt}} + f_{\text{HO,inf}} < Y_{\text{HO}} + S_{\text{consumed}}$, substitution of Eq. 10 to Eq. 8 gives:

$$m = \frac{f_{\text{HO,Capt}} + f_{\text{HO,inf}}}{X_{\text{HO,ML}} \left( \theta_x/\delta \right)/(1 + \theta_{\text{HO}} + \theta_x)}$$

The immigration rate of a specific $i$th OTU is:

$$m_i = \frac{f_{\text{HO,Capt}} + f_{\text{HO,inf}}}{X_{\text{HO,ML}} \left( \frac{\theta_x}{\delta} \right)/(1 + \theta_{\text{HO}} + \theta_x)}$$

The proportion of the $i$th OTU in the biomass sample ($f_{\text{16S}}$) can be quantified from the 16S rRNA amplicon sequencing data. The biomass concentration of the $i$th OTU can be defined by:

$$X_{\text{16S}} = X_{\text{16S}} Y_{\text{16S}} f_{\text{16S}}$$

where $X_{\text{16S}}$ is the biomass concentration of the $i$th OTU; $X_{\text{16S}}$ is the total concentration of solids; $Y_{\text{16S}}$ is the mass of DNA per solids; $f_{\text{16S}}$ is the proportion of the $i$th genus in total 16S rRNA gene amplicon sequence reads.

Therefore, the immigration rate of a specific $i$th OTU is expressed using 16S rRNA amplicon sequencing data:

$$m_i = \frac{\theta_x}{\delta} \left( \frac{f_{\text{HO,Capt}} \cdot X_{\text{cap},16S} \cdot Y_{\text{16S}} \cdot f_{\text{16S}}}{X_{\text{ML},16S} \cdot Y_{\text{ML}},Y_{\text{16S}} f_{\text{16S},i} f_{\text{16S}} \cdot \delta} \right)$$

Note that the maximum value of $m_i$ is $1$, so the calculated $m_i$ values when greater than $1$ were set to $1$. When $m_i = 1$ (i.e., the growth rate $\mu_{\text{HO,inf}} = 0$):

$$b_{\text{HO,inf}} = \frac{X_{\text{cap},16S} \cdot Y_{\text{16S}} \cdot f_{\text{16S}} f_{\text{16S},i} f_{\text{16S}}}{X_{\text{ML},16S} \cdot Y_{\text{ML}},Y_{\text{16S}} f_{\text{16S},i} f_{\text{16S}} \cdot \delta}$$

The net growth rate equals the growth rate minus decay rate, therefore the net growth rate of the $i$th OTU is:

$$\mu_{\text{HO,inf}} = \frac{X_{\text{cap},16S} \cdot Y_{\text{16S}} \cdot f_{\text{16S}} f_{\text{16S},i} f_{\text{16S}}}{X_{\text{ML},16S} \cdot Y_{\text{ML}},Y_{\text{16S}} f_{\text{16S},i} f_{\text{16S}} \cdot \delta}$$

The OTUs then can be classified into three groups:

(i) not in influent but in reactor community

(ii) growth group (calculated net growth rate $b_{\text{HO,inf}} > 0$)

(iii) complete immigration group (calculated net growth rate $b_{\text{HO,inf}} \leq 0$).

### DATA AVAILABILITY

The raw sequence data of 16S rRNA gene amplicons can be accessed at GenBank (Bio Project Accession number PRJNA580466).

Received: 23 September 2020; Accepted: 12 January 2021; Published online: 22 March 2021

### REFERENCES

1. Offerlu, I. D. et al. Combined niche and neutral effects in a microbial wastewater treatment community. Proc. Natl Acad. Sci. USA 107, 15345–15350 (2010).

2. Sloan, W. T. et al. Quantifying the roles of immigration and chance in shaping prokaryote community structure. Environ. Microbiol. 8, 732–740 (2006).

3. Burns, A. R. et al. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. ISME J. 10, 655–664 (2016).

4. Yuan, H., Mei, R., Liao, J. & Liu, W. T. Nexus of stochastic and deterministic processes on microbial community assembly in biological systems. Front. Microbiol. 10, 1536 (2019).

5. Mei, R., Kim, J., Wilson, F. P., Bocher, B. T. W. & Liu, W. T. Coupling growth kinetics modeling with machine learning reveals microbial immigration impacts and identifies key environmental parameters in a biological wastewater treatment process. Microbiome 7, 65 (2019).

6. Ning, D., Deng, Y., Tiedje, J. M. & Zhou, J. A general framework for quantitatively assessing ecological stochasticity. Proc. Natl Acad. Sci. USA 116, 16892–16898 (2019).

7. Saunders, A. M., Albertsen, M., Vollertsen, J. & Nielsen, P. H. The activated sludge ecosystem contains a core community of abundant organisms. ISME J. 10, 11–20 (2016).

8. Mei, R. & Liu, W. T. Quantifying the contribution of microbial immigration in engineered water systems. Microbiome 7, 144 (2019).

9. Vuono, D. C., Munakata-Marr, J., Spear, J. R. & Dreves, J. E. Disturbance opens recruitment sites for bacterial colonization in activated sludge. Environ. Microbiol. 18, 87–99 (2016).

10. Frigon, D. & Wells, G. Microbial immigration in wastewater treatment systems: analytical considerations and process implications. Curr. Opin. Biotechnol. 57, 151–159 (2019).

11. Guo, B. Cellular Metabolic Markers and Growth Dynamics Definition of Functional Groups in Activated Sludge Wastewater Treatment Heterotrophic Population. Doctor of Philosophy thesis, McGill Univ. (2019).

12. Sun, X., Sheng, Z. & Liu, Y. Effects of silver nanoparticles on microbial community structure in activated sludge. Sci. Total Environ. 443, 828–835 (2013).

13. Ali, M. et al. Importance of species sorting and immigration on the bacterial assembly of different-sized aggregates in a full-scale aerobic granular sludge plant. Environ. Sci. Technol. 53, 8291–8301 (2019).

14. Perez, M.V., Guerrero, L.D., Orellana, E., Figuerola, E.L. & Erijman, L. Time series genome-centric analysis unveils bacterial response to operational disturbance in activated sludge. mSystems 4(4), e00169–19, https://doi.org/10.1128/mSystems.00169-19 (2019).

15. 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).

16. Laurenti, M. et al. Bacterial segregation between biofilm and flocs improves the control of nitrite-oxidizing bacteria in mainstream partial nitritation and anammox processes. Water Res. 154, 104–116 (2019).

17. Wells, G. F. et al. Microbial biogeography across a full-scale wastewater treatment plant transect: evidence for immigration between coupled processes. Appl. Microbiol. Biotechnol. 98, 4723–4736 (2014).

18. Klappenbach, J. A., Dunbar, J. M. & Schmidt, T. M. rRNA operon copy number reflects ecological strategies of bacteria. Appl. Environ. Microbiol. 66, 1328–1333 (2000).
19. Nemergut, D. R. et al. Decreases in average bacterial community rRNA operon copy number during succession. *ISME J.* **10**, 1147–1156 (2016).

20. Guo, B., Liu, C., Gibson, C. & Frigon, D. Wastewater microbial community structure and functional traits change over short timescales. *Sci. Total Environ.* **662**, 779–785 (2019).

21. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).

22. Apprill, A., McNally, S., Parsons, R. & Weber, L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* **75**, 129–137 (2015).

23. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

24. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).

25. McDonald, D. et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618 (2012).

26. Werner, J. J. et al. Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *ISME J.* **6**, 94–103 (2012).

27. Oksanen, F. J. et al. vegan: Community Ecology Package. R package Version 2.4-3. https://CRAN.R-project.org/package=vegan (2017).

28. Langille, M. G. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).

29. Markowitz, V. M. et al. IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res.* **40**, D115–D122 (2012).

30. Grady, C. P. L. & Grady, C. P. L. *Biological Wastewater Treatment*. 3rd edn. (Taylor & Francis, 2011).

ACKNOWLEDGEMENTS
This work was funded by the Canada Research Chair (CRC) in Future Water Services (Y.L.) and Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (Y.L.).

AUTHOR CONTRIBUTIONS
B.G. collected and analyzed the samples, carried out data analysis, modeling and wrote the manuscript. Z.S. analyzed the samples and carried out analysis. Y.L. conceived and guided the study and revised the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41545-021-00112-7.

Correspondence and requests for materials should be addressed to Y.L.

Reprints and permission information is available at http://www.nature.com/reprints

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) 2021