

REVIEW

Open Access

Quantifying the contribution of microbial immigration in engineered water systems



Ran Mei and Wen-Tso Liu*

Abstract

Immigration is a process that can influence the assembly of microbial communities in natural and engineered environments. However, it remains challenging to quantitatively evaluate the contribution of this process to the microbial diversity and function in the receiving ecosystems. Currently used methods, i.e., counting shared microbial species, microbial source tracking, and neutral community model, rely on abundance profile to reveal the extent of overlapping between the upstream and downstream communities. Thus, they cannot suggest the quantitative contribution of immigrants to the downstream community function because activities of individual immigrants are not considered after entering the receiving environment. This limitation can be overcome by using an approach that couples a mass balance model with high-throughput DNA sequencing, i.e., ecogenomics-based mass balance. It calculates the net growth rate of individual microbial immigrants and partitions the entire community into active populations that contribute to the community function and inactive ones that carry minimal function. Linking activities of immigrants to their abundance further provides quantification of the contribution from an upstream environment to the downstream community. Considering only active populations can improve the accuracy of identifying key environmental parameters dictating process performance using methods such as machine learning.

Keywords: Microbial immigration, Mass balance, Engineered water systems, Microbiome

Introduction

Microbial communities play essential roles in biogeochemical cycles in natural and engineered ecosystems [1]. To study how different microorganisms assemble into a community and contribute to the function of an ecosystem, various mechanisms including the niche and neutral theories have been developed [2]. In the neutral theory of biodiversity and biogeography, immigration is one of the key stochastic processes that change the community assemblage together with death and birth [3]. This process, sometimes referred to as migration [4], is originally used in macroecology to estimate the rate of new bird species entering a remote island from the nearest land mass, i.e., the chance of immigration, which plays a pivotal role in the equilibrium of island fauna's diversity [5]. As the definition of immigration can vary considerably [6], this review adopts the one stated by Bell [4], and defines immigration as the process of a microbial individual being added to a local community

from the species pool of the metacommunity, which consists of a set of local communities that are physically linked by immigration and can exchange colonists of multiple species [7]. A similar term often used is dispersal [8–10]. While dispersal and immigration may slightly differ in specific context and one may even include the other [6], there is still no consensus on the difference [10–13]. Therefore, we do not attempt to discuss the differences between immigration and dispersal here. We further consider microorganisms that arrive at the local communities all as immigrants, regardless of how they arrive (e.g., facilitated by cell motility or flow of water or air) and how they contribute to the local community after arrival.

In natural microbial ecosystems, immigration from an upstream community to a downstream community can be exemplified by microbial immigration from Africa dust to European aquatic environments [14], from running waters to forest lakes [15], and from river water to estuarine and offshore environments [16]. Microbial immigration also widely occurs in engineered environments, and this movement often occurs in a more

* Correspondence: wliu@illinois.edu
Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA



controlled and directed manner than in natural environments. In drinking water production systems (Fig. 1a), microorganisms in the source water, e.g., groundwater and surface water, can serve as inocula to seed the communities along the treatment process [17]. Next, microbes growing on the treatment units, e.g., surfaces of reactor walls or filtration media, can be released to subsequent processes [18] and then to the distribution system depending on whether disinfection is practiced [19] or not [20]. In the distribution network and indoor plumbing, bidirectional immigration can occur between the water phase and the biofilms on the inner pipe surface [21], which can be further elevated through water stagnation [22]. A wastewater treatment plant (WWTP) is another good example of microbial immigration (Fig. 1b), because different bioreactors are physically connected, and the flux of biomass can be higher than those in natural systems. In WWTPs, microbes present in raw sewer [23] or nitrifying trickling filter [24] were reported to have impact on the downstream activated sludge communities. The activated sludge biomass can also act as a source of immigration to the downstream anaerobic digester microbiome [25]. The WWTP effluent can impact the receiving water body community [26], elevating the abundance of human gut-related microbes and antibiotic resistance genes [27].

While microbial immigration is frequently reported in engineered water systems, it remains challenging to quantitatively address “to what extent immigration contributes to the assembly and function of the downstream

community?” In this article, we focus on methodology quantifying microbial immigration by first reviewing methods that are currently used and identifying their limitations. Then, an approach that calculates the net growth rates of individual microbial immigrants in the downstream community is reviewed. It couples a mass balance model with high-throughput DNA sequencing to partition microbial assembly into active populations that contribute to community function and inactive ones that carry minimal microbial function. Its potential use together with machine learning to identify key environmental parameters affecting the microbial ecosystem’s function is discussed.

Methods commonly used to evaluate immigration impact

Figure 2 illustrates three methods commonly used to evaluate immigration impact in microbial ecosystems. The first approach simply counts shared species between the upstream and downstream communities, as visualized by a Venn diagram (Fig. 2a). For example, McLellan et al. reported that untreated sewage communities shared more species with surface water communities than with human fecal sources, indicating surface water such as rainwater and storm water can modulate sewage microbial community composition [28]. Lee et al. compared microbial communities between influent wastewater and the downstream activated sludge in four full-scale WWTPs and reported that shared species accounted for 12.2%, 7.5%, 15.2%, and 7.6% of total sequences in activated sludge, respectively

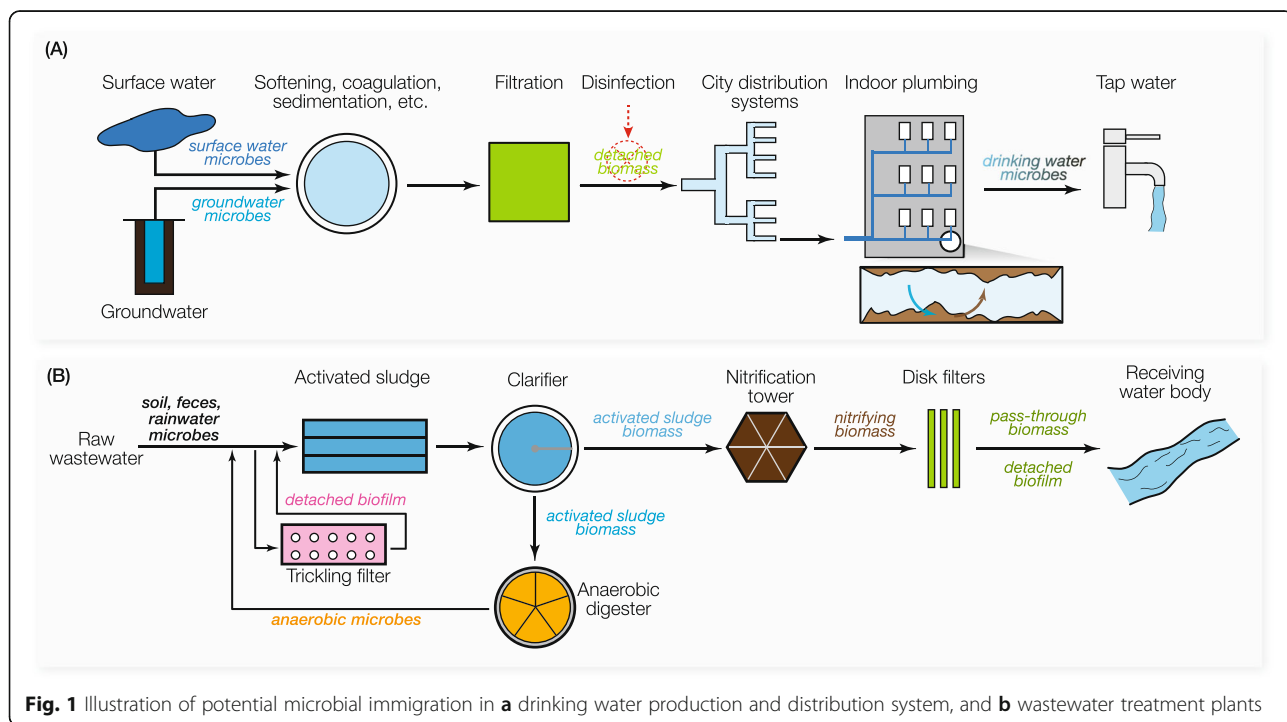
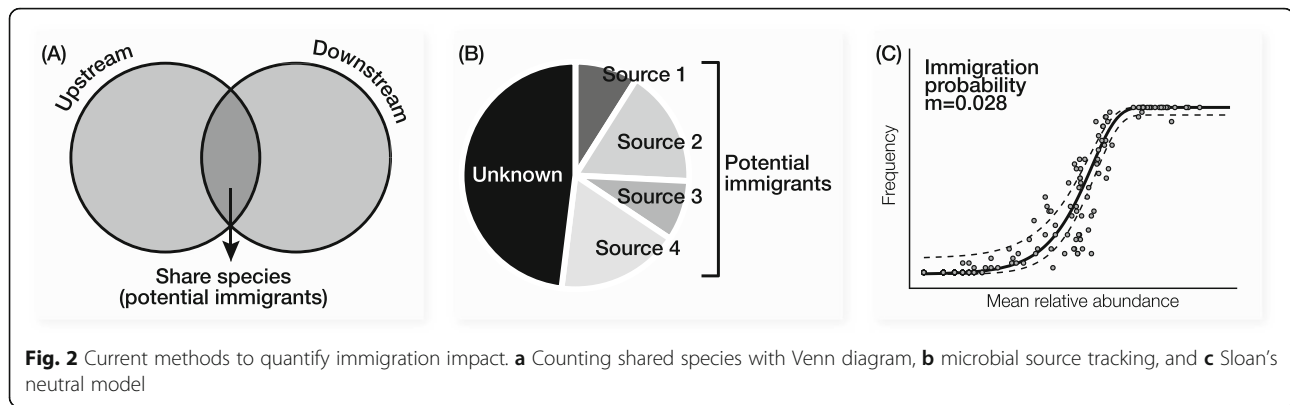


Fig. 1 Illustration of potential microbial immigration in **a** drinking water production and distribution system, and **b** wastewater treatment plants



[23]. As the authors stated, these numbers only implied “contribution from influent wastewater communities to some extent”, while the exact contribution to the function of each activated sludge system cannot be quantified. This uncertainty is due to the fact that the microbial activity of immigrants cannot be determined. Shared species can be generalists that function in both upstream and downstream environments or be detectable but remain dormant in activated sludge due to their high abundance in the influent and low adaptability (i.e., low activity) to the new environmental conditions.

The second approach is microbial source tracking that estimates the proportion of taxa in the downstream or sink community coming from multiple upstream or source environments [29] (Fig. 2b). The basic rationale is that more abundant taxa in the source have higher probabilities to be observed in the sink, which represents the contribution of each source. This method has been applied to study a sink environment that receives immigration from multiple sources, such as residential kitchen microbiome subject to source microbiota of the human palm skin, produces, and faucet water [30] and public restroom microbiome subject to source microbiota of the soil, water, human urine, gut, mouth, and skin [31]. However, the method assumes that all the observed microbial populations in the sink community come from source environments and ignores the fact that some active microorganisms can undergo rapid reproduction after entering the sink. When the abundance of immigrants increases, the source tracking method cannot fully explain this fraction of community composition from known sources. This limitation leads to the observation that sometimes the majority of the sink community is labeled as unknown, whereas only a small proportion can be explained by known sources. The observation of unknown sources is especially common in systems with high microbial activities, such as wastewater treatment processes [32–34].

The third approach uses the neutral model (Fig. 2c) developed by Sloan and coworkers [35]. By determining the abundance-frequency distribution of individual

microbial species, a species-independent immigration probability m is calculated, which is uniform for every community member. A small m value suggests that the community as a whole is comprised of a low proportion of immigrants from the source. This model has been frequently used to evaluate the relative importance of neutral mechanism and directly assess the immigration rate at community level by calculating the m value. Using this model, Ayarza and Erijman revealed that neutral process was important in the assembly of activated sludge community [36]. In a drinking water distribution system, the model was used to demonstrate that the role of immigration from city water supply to tap water was higher at the proximal end than at the distal end of indoor plumbing [22]. Likewise, studies have compared the m values to reveal higher immigration impact in planktonic communities than in sedimentary communities in Yangtze River [37], and in deep-water communities than in surface water communities [38]. Despite the success of the model in explaining the general trend of abundance-frequency distribution, there are always species significantly deviating from the S-shape fitting curve. This is likely caused by assuming a constant immigration rate for all community members. It ignores the fact that immigrants with active microbial growth in the new environments can become more abundant than the prediction.

All the three methods described above merely enumerate the number of immigrants and cannot fully reveal the impact of microbial immigration on community functions. As microorganisms can carry considerably diverse activities after entering a new environment, in contrast to inert and homogeneous particles, it is important to address “how many microbial immigrants are able to actively contribute to ecological functions.”

Quantifying immigration impact with the consideration of microbial activity

To quantitatively evaluate the immigration impact, both the abundance and activity of individual immigrants should

be considered. Compared to determining abundance, assessing the activities or growth of individual immigrants in a given microbial ecosystem is challenging. In pure cultures, microbial activity can be assessed by measuring substrate consumption, metabolite production, or cell density change during a period of incubation. However, only a small fraction of microbes in nature can be cultivated and microbial activities determined in pure culture can differ drastically in a complex community under environmental conditions [39]. Sequencing 16S ribosomal RNA (rRNA) genes and other biomarkers is often used to identify microorganisms present in the environment but cannot effectively distinguish active populations from inactive or dormant species. Likewise, metagenomics reveals functional potentials of community members but cannot discern expressed and non-expressed pathways [40]. Directly sequencing rRNA can identify active microbes, but the consistency of this approach can be affected by the differences between rRNA content and microbial activity [41]. Sequencing messenger RNA (mRNA), i.e., metatranscriptomics, provides accurate identification of highly expressed genes and active populations from environmental samples [42], but this approach is challenged by the scarcity of well-annotated high-quality reference genomes [43]. Nucleotide sequencing can be coupled with methods such as microautoradiography [44], stable isotope probing [45], or nano secondary ion mass spectrometry [46] that label specific substrates to link substrates uptake activity with microbial identity. It is also possible to label and visualize specific active populations using fluorescence in situ hybridization designed to target rRNA [47]. These methods however cannot target all community members due to the cost and time associated with labeling individual substrates or organisms. In addition, metaproteomics and metabolomics can characterize the entire collection of proteins or metabolites of a given sample, and provide direct measurement of microbial activity, but also face challenges on preparing high-quality samples from complex environments and on linking proteins/metabolites with microbial identity [40]. Overall, it is still expensive and time-consuming using these ecological tools to quantify the in situ activities of most microbial populations in a complex ecosystem.

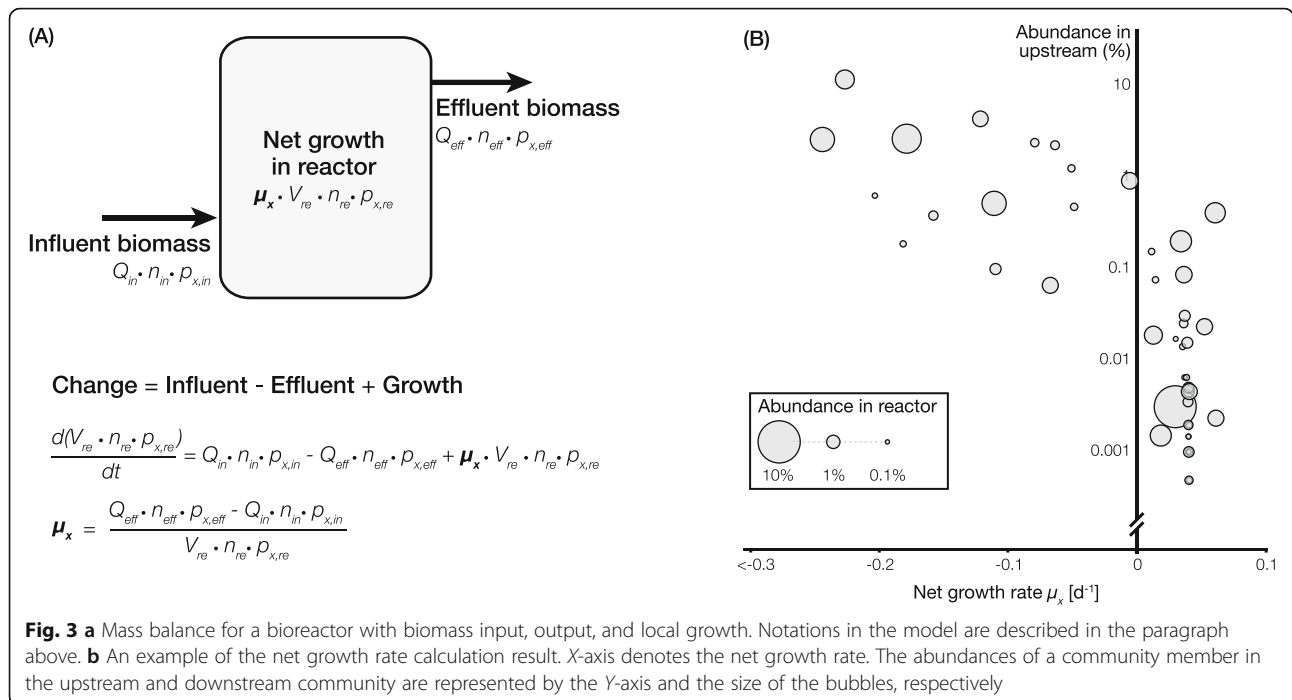
Quantifying immigration impact at individual population level using ecogenomics-based mass balance approach

Recent studies [48–51] have presented a high-throughput and quantitative way to assess the immigration impact on an open engineered water system by combining a mass balance concept with ecogenomics tools, i.e., the ecogenomics-based mass balance. Using a biological reactor treating wastewater as an example (Fig. 3a), it receives influent biomass discharged from the upstream system with a flow rate Q_{in} and unit cell number n_{in} . The

relative abundance of any given microorganism x in the influent stream ($p_{x, in}$) can be determined by 16S rRNA gene sequencing. Thus, the absolute number of x in the influent can be described as $Q_{in} \cdot n_{in} \cdot p_{x, in}$ if Q_{in} and n_{in} are properly measured. Similarly, the number of x in the effluent can be described similarly as $Q_{eff} \cdot n_{eff} \cdot p_{x, eff}$ and multiple influent/effluent streams are allowed. The growth of x in the reactor is expressed with a first order reaction ($\mu_x \cdot V_{re} \cdot n_{re} \cdot p_{x, re}$), i.e., the product of a net growth rate constant μ_x and the absolute number of x in the reactor $V_{re} \cdot n_{re} \cdot p_{x, re}$. When the reactor is operated at a steady state and the number of x in the reactor does not change through time, the sum of x in the influent, effluent, and its growth is zero, which allows the calculation of μ_x . This growth rate constant is specific for each community member, which can be positive (i.e., net growth and active ecological function) or negative (i.e., net decay and minimal ecological function). When the growth rates of all community members are plotted against their abundance, one can easily tell whether the abundant populations in the upstream environment (major immigrants) are active or not in the downstream community. In the example illustrated in Fig. 3b, those located at the upper part of the figure tend to have negative growth rates, suggesting immigrants from the upstream exhibit low activity, although they still retain high abundance as indicated by the large bubble size. In contrast, active populations, those with positive growth rate, in the reactor are not derived from the upstream process. Overall, although immigrants represent a large proportion of observed abundance in this case, their contribution to the community function is likely low considering their negative growth rates.

Saunders et al. first developed this method and studied three WWTPs [48]. Thirty-five percent of the observed species in activated sludge reactors were also detected in the influent wastewater, suggesting a strong immigration impact. However, the mass balance revealed that majority of the shared species had negative net growth rate, suggesting they did not actively contribute to the metabolisms in activated sludge. There were a few immigrants with positive growth rate in activated sludge, indicating that they were indeed active in situ. Overall, the authors concluded a modest impact of immigration on the activated sludge community, considering there were both inactive and active immigrants with moderate abundance.

Mei et al. applied the ecogenomics-based mass balance approach to anaerobic digesters that receive massive biomass from upstream activated sludge [49]. Based on the result that populations with negative net growth rate accounted for 25% of total sequences in digesters, a strong immigration impact with the feed aerobic wasted sludge was concluded. Phylogenetic analysis confirmed that inactive populations were associated with aerobes or facultative anaerobes, whereas active populations were



associated with obligate anaerobes. This study also reported the bias associated with the use of 16S rRNA-based relative abundance as an activity indicator in environments under high immigration impact. It is possible that some immigrants, which are the major populations in the previous environment, can still contain high copy number of rRNA after moving to a downstream environment. Thus, rRNA-based calculation can overestimate the relative abundance and contribution of these populations in the downstream community [49]. Such rationale was further used to multiple full-scale digesters around the world [52, 53], and the findings revealed that immigration from feed sludge to the digester communities was a ubiquitous phenomenon and the extent of contribution was also influenced by the operation conditions and pretreatments related to the anaerobic digestion.

The ecogenomics-based mass balance approach was also applied to an industrial WWTP to demonstrate its effectiveness in teasing apart the interaction between neutral and niche-based mechanisms [50]. In the studied WWTP, immigrants from an upstream anaerobic reactor were inactive (net growth rate ≤ 0) and represented a negligible fraction (1% of the total sequences) in the downstream activated sludge community, implying a weak immigration impact. But these immigrants were found to affect the prediction of key environmental parameters from community composition using a machine-learning tool [54]. To do so, a supervised learning regressor was first trained on a set of samples with known physiochemical parameters such as temperature, pH, and nutrient concentrations and then used to

predict the target values of the remaining samples. Parameters with higher prediction accuracy played more important roles in shaping the microbial community. After removing inactive immigrants from the downstream community, the prediction accuracy greatly improved. This result suggests that more cautious interpretation should be made to identify the key environmental parameters based on community composition. Commonly used methods, including k-means clustering [55], principal components analysis [56], principal coordinate analysis [57], non-metric multidimensional scaling [58], and redundancy analysis [59], solely rely on DNA-based microbial abundance, but pay little attention to the existence of inactive immigrants. The efforts of correlating observed species abundance with environmental conditions would be notably biased in an open ecosystem where inactive populations are introduced by immigration.

Wastewater systems are often designed with complex process configuration where high biomass flux from one process to the next can take place. Sampling and controlling in these systems at different temporal and spatial scales are easier than natural environments. Therefore, these environments present an excellent opportunity to apply the ecogenomics-based mass balance approach to quantify the contribution of microbial immigration to community composition and function. Furthermore, this method can be applied to non-wastewater systems where microbial immigration is commonly present but the contribution is rarely quantified. The differentiation of inactive populations is specifically valuable in those

environments where the growth rates of microorganisms can be more heterogeneous than in highly selective wastewater systems. For example, biofilm growth in the drinking water distribution pipe has been recognized as an important process that affects drinking water quality [21, 22]. The mass balance model can be used to characterize the growth and immigration of different organisms in the biofilm, especially those posing risks to human health. To do so, a section of pipe can be considered as the control volume, with fresh city water and tap water as the influent and effluent, respectively. Organisms that can scavenge substrates in the pipe will exhibit higher growth rates and can be released into the tap water. Different conditions and parameters related to the mass balance can be tested to assess their impacts on immigration. Some of them include the disinfection methods of the city water supply, the sizes and materials of the pipe, the temperature of the environment, and the period of the water stagnation. These results can provide guidance to improve drinking water quality and prevent waterborne disease outbreak from the aspect of microbial ecology. While the measurements of mass flux and cell count related to biofilm can be challenging, they can be solved for example by harvesting the biofilm after a period of development and by enumerating cell number with flow cytometry as demonstrated in a recent study [51]. Optical coherence tomography is another effective and non-destructive way to determine the biofilm mass and possibly mass change (e.g., in term of volume) on the pipe inner surface [60]. In other systems where diverse ecological functions are carried out, functional genes, such as *mcrA* for methanogenesis and *amoA* for nitrification, can be used to monitor a subset of populations with specific function(s) instead of 16S rRNA gene. Besides marker genes, metagenomics and metatranscriptomics can be used to estimate the abundance and activity, respectively, of individual immigrants with higher resolution. In addition, the immigration of viruses [61] and eukaryotes [62] from an upstream process to a downstream process can be monitored and quantified, in addition to prokaryotic populations.

The ecogenomics-based mass balance approach is effective in differentiating active and inactive populations resulted from immigration than the three methods discussed earlier that do not consider immigrants' activities (Table 1). It evaluates immigration impacts by considering the abundance and activities of all immigrants. The net growth rate is calculated in a high-throughput manner and is specific for individual immigrants, unlike an index derived for a whole community from other three methods. The method can be further applied to complex ecosystems by including multiple influent and effluent streams in the calculation. Still, the mass balance approach should be used carefully with good experimental

Table 1 Comparison of the commonly used methods that quantify immigration impact and the ecogenomics-based mass balance

Method	Venn diagram	Source tracking	Neutral model	Ecogenomics-based mass balance
Abundance of total immigrants	✓	✓	✓	✓
Abundance of individual immigrants	X	X	X	✓
Activity of individual immigrants	X	X	X	✓
Multiple upstream environments	✓	✓	X	✓
Multiple downstream environments	✓	X	X	✓
Cell number estimation	X	X	X	✓
Flux measurement	X	X	X	✓
Steady-state assumption	X	X	X	✓
Repeated sampling	X	✓	✓	✓

design. It assumes that the system is in a steady state and requires multiple measurements of cell number and biomass flux within a properly defined control volume. When 16S rRNA gene is used as the biomarker, biases associated with DNA extraction, PCR amplification, rRNA copy number, and amplicon sequencing are still present [63]. The calculated net growth rate is only a proxy of the in situ activity and cannot perfectly reflect the behaviors of individual community members. It remains challenging to apply the approach to systems with attached microbial growth such as granular sludge bioreactors, due to the difficulty in estimating total biomass and biomass flux, as well as the heterogeneity in microbial compositions within the consortia. Thus, more research efforts are needed in this direction.

Conclusion

Microbial immigration is a ubiquitous and important process occurring in engineered water systems, and it allows microbes present in an upstream system to influence the microbial assembly and function in a downstream receiving system after entering. To understand the impact of microbial immigration, qualitative and quantitative methods are necessary and have been developed. Commonly used methods are recognized to have limitations in quantifying the immigration impacts. The ecogenomics-based mass balance approach provides a solution by quantitatively determining the activity profile of all microbial populations in a community. This approach can effectively identify inactive populations, especially those resulted from immigration, and pinpoint microorganisms that are actually carrying out the

process of interest. Furthermore, when coupled with methods such as machine learning, it can better identify key environmental parameters affecting system performance, which can guide the monitoring and designing of biological processes. It is foreseen that such an approach can be widely applied to various engineered and possibly natural environments, where the contribution of microbial immigration remains to be further characterized.

Abbreviations

mRNA: Messenger RNA; rRNA: Ribosomal RNA; WWTP: Wastewater treatment plant

Acknowledgements

Not applicable.

Authors' contributions

RM conceived the idea and wrote the manuscript. WTL conceived the idea and reviewed the manuscript. Both authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 23 April 2019 Accepted: 17 October 2019

Published online: 06 November 2019

References

- Cira NJ, Pearce MT, Quake SR. Neutral and selective dynamics in a synthetic microbial community. *Proc Natl Acad Sci*. 2018;115(42):9842–8.
- Fargione J, Brown CS, Tilman D. Community assembly and invasion: an experimental test of neutral versus niche processes. *Proc Natl Acad Sci*. 2003;100(15):8916–20.
- Hubbell SP. The unified neutral theory of biodiversity and biogeography (MPB-32). Princeton: Princeton University Press; 2001.
- Bell G. The distribution of abundance in neutral communities. *Am Nat*. 2000;155(5):606–17.
- MacArthur RH, Wilson EO. An equilibrium theory of insular zoogeography. *Evolution*. 1963;17(4):373–87.
- Schneider F. Dispersal and migration. *Annu Rev Entomol*. 1962;7(1):223–42.
- Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, Holt RD, Shurin JB, Law R, Tilman D, et al. The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett*. 2004;7(7):601–13.
- Vellend BM. Conceptual synthesis in community ecology. *Q Rev Biol*. 2010;85(2):183–206.
- Konopka A, Lindemann S, Fredrickson J. Dynamics in microbial communities: unraveling mechanisms to identify principles. *ISME J*. 2015;9(7):1488–95.
- Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, Rockhold ML, Konopka A. Quantifying community assembly processes and identifying features that impose them. *ISME J*. 2013;7(11):2069–79.
- Fodelianakis S, Lorz A, Valenzuela-Cuevas A, Barozzi A, Booth JM, Daffonchio D. Dispersal homogenizes communities via immigration even at low rates in a simplified synthetic bacterial metacommunity. *Nat Commun*. 2019;10(1):1314.
- Frigon D, Wells G. Microbial immigration in wastewater treatment systems: analytical considerations and process implications. *Curr Opin Biotechnol*. 2019;57:151–9.
- Gilpin ME, Diamond JM. Calculation of immigration and extinction curves from the species-area-distance relation. *Proc Natl Acad Sci*. 1976;73(11):4130–4.
- Hervás A, Camarero L, Reche I, Casamayor EO. Viability and potential for immigration of airborne bacteria from Africa that reach high mountain lakes in Europe. *Environ Microbiol*. 2009;11(6):1612–23.
- Lindström ES, Bergström A-K. Influence of inlet bacteria on bacterioplankton assemblage composition in lakes of different hydraulic retention time. *Limnol Oceanogr*. 2004;49(1):125–36.
- Kisand V, Andersson N, Wikner J. Bacterial freshwater species successfully immigrate to the brackish water environment in the northern Baltic. *Limnol Oceanogr*. 2005;50(3):945–56.
- Pinto AJ, Xi C, Raskin L. Bacterial community structure in the drinking water microbiome is governed by filtration processes. *Environ Sci Technol*. 2012;46(16):8851–9.
- Kotlarz N, Rockey N, Olson TM, Haig S-J, Sanford L, LiPuma JJ, Raskin L. Biofilms in full-scale drinking water ozone contactors contribute viable bacteria to ozonated water. *Environ Sci Technol*. 2018;52(5):2618–28.
- Zhang Y, Oh S, Liu WT. Impact of drinking water treatment and distribution on the microbiome continuum: an ecological disturbance's perspective. *Environ Microbiol*. 2017;19(8):3163–74.
- Lautenschlager K, Hwang C, Ling F, Liu W-T, Boon N, Köster O, Egli T, Hammes F. Abundance and composition of indigenous bacterial communities in a multi-step biofiltration-based drinking water treatment plant. *Water Res*. 2014;62:40–52.
- Ling F, Hwang C, LeChevallier MW, Andersen GL, Liu W-T. Core-satellite populations and seasonality of water meter biofilms in a metropolitan drinking water distribution system. *ISME J*. 2015;10(3):582–95.
- Ling F, Whitaker R, LeChevallier MW, Liu W-T. Drinking water microbiome assembly induced by water stagnation. *ISME J*. 2018;12(6):1520–31.
- Lee S-H, Kang H-J, Park H-D. Influence of influent wastewater communities on temporal variation of activated sludge communities. *Water Res*. 2015;73:132–44.
- Wells GF, Wu CH, Piceno YM, Eggleston B, Brodie EL, DeSantis TZ, Andersen GL, Hazen TC, Francis CA, Criddle CS. Microbial biogeography across a full-scale wastewater treatment plant transect: evidence for immigration between coupled processes. *Appl Microbiol Biotechnol*. 2014;98(10):4723–36.
- Ju F, Lau F, Zhang T. Linking microbial community, environmental variables, and methanogenesis in anaerobic biogas digesters of chemically enhanced primary treatment sludge. *Environ Sci Technol*. 2017;51(7):3982–92.
- Price JR, Ledford SH, Ryan MO, Toran L, Sales CM. Wastewater treatment plant effluent introduces recoverable shifts in microbial community composition in receiving streams. *Sci Total Environ*. 2018;613:1104–16.
- Zhang S-Y, Tsementzi D, Hatt JK, Bivins A, Khelurkar N, Brown J, Tripathi SN, Konstantinidis KT. Intensive allochthonous inputs along the Ganges River and their effect on microbial community composition and dynamics. *Environ Microbiol*. 2018;21(1):182–96.
- McLellan S, Huse S, Mueller-Spitz S, Andreishcheva E, Sogin M. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environ Microbiol*. 2010;12(2):378–92.
- Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R, Kelley ST. Bayesian community-wide culture-independent microbial source tracking. *Nat Methods*. 2011;8:761.
- Flores GE, Bates ST, Caporaso JG, Lauber CL, Leff JW, Knight R, Fierer N. Diversity, distribution and sources of bacteria in residential kitchens. *Environ Microbiol*. 2013;15(2):588–96.
- Flores GE, Bates ST, Knights D, Lauber CL, Stombaugh J, Knight R, Fierer N. Microbial biogeography of public restroom surfaces. *PLoS One*. 2011;6(11):e28132.
- Shanks OC, Newton RJ, Kelty CA, Huse SM, Sogin ML, McLellan SL. Comparison of the microbial community structures of untreated wastewaters from different geographic locales. *Appl Environ Microbiol*. 2013;79(9):2906.
- Staley C, Gould TJ, Wang P, Phillips J, Cotner JB, Sadowsky MJ. Species sorting and seasonal dynamics primarily shape bacterial communities in the Upper Mississippi River. *Sci Total Environ*. 2015;505:435–45.
- Ahmed W, Staley C, Sadowsky MJ, Gyawali P, Sidhu JPS, Palmer A, Beale DJ, Toze S. Toolbox approaches using molecular markers and 16S rRNA gene

- amplicon data sets for identification of fecal pollution in surface water. *Appl Environ Microbiol.* 2015;81(20):7067.
35. Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol.* 2006;8(4):732–40.
 36. Ayarza JM, Erijman L. Balance of neutral and deterministic components in the dynamics of activated sludge floc assembly. *Microb Ecol.* 2011; 61(3):486–95.
 37. Liu T, Zhang AN, Wang J, Liu S, Jiang X, Dang C, Ma T, Liu S, Chen Q, Xie S. Integrated biogeography of planktonic and sedimentary bacterial communities in the Yangtze River. *Microbiome.* 2018;6(1):16.
 38. Logares R, Lindström ES, Langenheder S, Logue JB, Paterson H, Laybourn-Parry J, Rengefors K, Tranvik L, Bertilsson S. Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME J.* 2013;7(5):937.
 39. Price PB, Sowers T. Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *Proc Natl Acad Sci U S A.* 2004;101(13):4631–6.
 40. Su C, Lei L, Duan Y, Zhang K-Q, Yang J. Culture-independent methods for studying environmental microorganisms: methods, application, and perspective. *Appl Microbiol Biotechnol.* 2012;93(3):993–1003.
 41. Blazewicz SJ, Barnard RL, Daly RA, Firestone MK. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J.* 2013;7(11):2061–8.
 42. Poretsky RS, Gifford S, Rinta-Kanto J, Vila-Costa M, Moran MA. Analyzing gene expression from marine microbial communities using environmental transcriptomics. *J Vis Exp.* 2009;24:e1086.
 43. Vanwonderghem I, Jensen PD, Ho DP, Batstone DJ, Tyson GW. Linking microbial community structure, interactions and function in anaerobic digesters using new molecular techniques. *Curr Opin Biotechnol.* 2014; 27:55–64.
 44. Ito T, Yoshiguchi K, Ariesyady HD, Okabe S. Identification of a novel acetate-utilizing bacterium belonging to Synergistes group 4 in anaerobic digester sludge. *ISME J.* 2011;5(12):1844–56.
 45. Hatamoto M, Imachi H, Yashiro Y, Ohashi A, Harada H. Detection of active butyrate-degrading microorganisms in methanogenic sludges by RNA-based stable isotope probing. *Appl Environ Microbiol.* 2008;74(11):3610–4.
 46. Li T, Wu TD, Mazéas L, Toffin L, Guerquin-Kern JL, Leblon G, Bouchez T. Simultaneous analysis of microbial identity and function using NanoSIMS. *Environ Microbiol.* 2008;10(3):580–8.
 47. Sekiguchi Y, Kamagata Y, Nakamura K, Ohashi A, Harada H. Fluorescence in situ hybridization using 16S rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. *Appl Environ Microbiol.* 1999;65(3):1280–8.
 48. Saunders AM, Albertsen M, Vollertsen J, Nielsen PH. The activated sludge ecosystem contains a core community of abundant organisms. *ISME J.* 2015;10(1):11–20.
 49. Mei R, Narihiro T, Nobu MK, Kuroda K, Liu WT. Evaluating digestion efficiency in full-scale anaerobic digesters by identifying active microbial populations through the lens of microbial activity. *Sci Rep.* 2016;6:34090.
 50. Mei R, Kim J, Wilson FP, Bocher BTW, Liu WT. Coupling growth kinetics modeling with machine learning reveals microbial immigration impacts and identifies key environmental parameters in a biological wastewater treatment process. *Microbiome.* 2019;7(1):65.
 51. Cheng H, Cheng D, Mao J, Lu T, Hong P-Y. Identification and characterization of core sludge and biofilm microbiota in anaerobic membrane bioreactors. *Environ Int.* 2019;133:105165.
 52. Mei R, Nobu MK, Narihiro T, Kuroda K, Munoz Sierra J, Wu Z, Ye L, Lee PKH, Lee PH, van Lier JB, et al. Operation-driven heterogeneity and overlooked feed-associated populations in global anaerobic digester microbiome. *Water Res.* 2017;124:77–84.
 53. Kirkegaard RH, McIlroy SJ, Kristensen JM, Nierychlo M, Karst SM, Dueholm MS, Albertsen M, Nielsen PH. The impact of immigration on microbial community composition in full-scale anaerobic digesters. *Sci Rep.* 2017;7(1):9343.
 54. Bokulich N, Dillon M, Bolyen E, Kaehler B, Huttley G, Caporaso J. q2-sample-classifier: machine-learning tools for microbiome classification and regression. *J Open Source Softw.* 2018;3:934.
 55. McHardy IH, Goudarzi M, Tong M, Ruegger PM, Schwager E, Weger JR, Graeber TG, Sonnenburg JL, Horvath S, Huttenhower C. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome.* 2013;1(1):17.
 56. Mei R, Narihiro T, Nobu MK, Liu WT. Effects of heat shocks on microbial community structure and microbial activity of a methanogenic enrichment degrading benzoate. *Lett Appl Microbiol.* 2016;63(5):356–62.
 57. Mei R, Nobu MK, Narihiro T, Yu J, Sathyagal A, Willman E, Liu WT. Novel *Geobacter* species and diverse methanogens contribute to enhanced methane production in media-added methanogenic reactors. *Water Res.* 2018;147(147):403–12.
 58. Ye L, Amberg J, Chapman D, Gaikowski M, Liu W-T. Fish gut microbiota analysis differentiates physiology and behavior of invasive Asian carp and indigenous American fish. *ISME J.* 2014;8(3):541–51.
 59. Narihiro T, Kim NK, Mei R, Nobu MK, Liu WT. Microbial community analysis of anaerobic reactors treating soft drink wastewater. *PLoS One.* 2015;10(3): e0119131.
 60. Shen Y, Huang PC, Huang C, Sun P, Monroy GL, Wu W, Lin J, Espinosa-Marzal RM, Boppart SA, Liu W-T, et al. Effect of divalent ions and a polyphosphate on composition, structure, and stiffness of simulated drinking water biofilms. *NPJ Biofilms Microbiomes.* 2018;4(1):15.
 61. Tamaki H, Zhang R, Angly FE, Nakamura S, Hong PY, Yasunaga T, Kamagata Y, Liu WT. Metagenomic analysis of DNA viruses in a wastewater treatment plant in tropical climate. *Environ Microbiol.* 2012;14(2):441–52.
 62. Hirakata Y, Hatamoto M, Oshiki M, Watari T, Kuroda K, Araki N, Yamaguchi T. Temporal variation of eukaryotic community structures in UASB reactor treating domestic sewage as revealed by 18S rRNA gene sequencing. *Sci Rep.* 2019;9(1):12783.
 63. Dohm JC, Lottaz C, Borodina T, Himmelbauer H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res.* 2008;36(16):e105.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

