Deterministic and stochastic processes influence the reproducibility of microbial community assembly during aerobic granulation

Raquel Liébana [a]*, Enikö Szabó [a], Frank Persson [a], Oskar Modin [a], Malte Hermansson [b] and Britt-Marie Wilén [a]

[a] Division of Water Environment Technology, Department of Architecture and Civil Engineering, Chalmers University of Technology, SE-41296 Gothenburg, Sweden

[b] Department of Chemistry and Molecular Biology, University of Gothenburg, SE-40530 Gothenburg, Sweden

*Corresponding author. Email address: raquel.liebana@chalmers.se

ABSTRACT

Aerobic granular sludge is a biofilm process which has received much attention during the last decades due to the advantages compared to conventional wastewater treatment. However, the ecological processes responsible for microbial community assembly during granulation are still poorly understood and little is known about the reproducibility in treatment performance and community structure. Here we show that both deterministic and stochastic processes exert a dynamic influence during different successional stages of microbial community assembly into granular sludge. During granulation, three replicate sequencing batch reactors followed a similar temporal successional trajectory in the most abundant taxa and similar diversity dynamics. Deterministic factors dominated the assembly of the most abundant community members as the microbial community transitioned from floccular to granular form. Stochastic factors mostly affected rare members of the communities. Stochasticity caused the microbial community structure to diverge in one of the reactors; however, this did not have an impact on the treatment performance. The results showed that the reactors functions and the dynamics of the most abundant community members are highly reproducible during the formation of aerobic granules.

Introduction

The structure of microbial communities depends on complex and dynamic ecological mechanisms. Studying the compositional differences between two different communities (β-diversity) is useful in order to understand the underlying ecological mechanisms that have shaped the communities, since it provides a link between biodiversity at the local scale (α-diversity) and at the regional scale (γ-diversity) (Anderson *et al.*, 2011, Shade, 2016). Vellend (2010) developed a conceptual framework, modified by Nemergut *et al.* (2013), where the ecological mechanisms responsible for community assembly are framed into four fundamental processes: selection, dispersal, diversification and drift. Selection results from deterministic factors that modify the community structure according to the environmental conditions, differences in fitness between individuals (e.g. survival, growth) and microbial interactions (e.g. predation, host selection, mutualisms, etc.). Dispersal results both from deterministic and stochastic factors which enable the movement and establishment of organisms among communities. Diversification results mainly from stochastic factors which generate genetic variation. Drift refers to stochastic changes due to birth, death and reproduction. These four processes act simultaneously in natural ecosystems and their influences vary in time and space (Nemergut *et al.*, 2013, Zhou & Ning, 2017).

Generally, the design and development of wastewater treatment systems is hindered by the lack of understanding of the ecology of the microbial communities on which they rely (McMahon *et al.*, 2007, Xia *et al.*, 2018). The importance ascribed to the different ecological mechanisms for shaping the microbial community composition in wastewater treatment bioreactors vary in the literature. Some studies report similar microbial communities in replicate reactors, such as membrane bioreactors (Falk *et al.*, 2009), anaerobic digesters (Vanwonterghem *et al.*, 2014, Lucas *et al.*, 2015, Luo *et al.*, 2015), microbial fuel cells (El-Chakhtoura *et al.*, 2014) and sequencing batch reactors (Wittebolle *et al.*, 2009, Ayarza *et al.*, 2010), due to selection provided by the reactor environmental conditions. On the contrary, diverging trajectories have been reported in replicate microbial electrolysis cell reactors (Zhou *et al.*, 2013), sequencing batch reactors (Kaewpipat & Grady, 2002) and anaerobic digesters (Solli *et al.*, 2014, Han *et al.*, 2016), due to the roles of stochastic factors. However, deterministic and stochastic factors are not mutually exclusive but exert a combined effect on the microbial community assembly and need to be assessed in conjunction (Nemergut *et al.*, 2013, Zhou & Ning, 2017).

Wastewater treatment bioreactors are engineered to select for several functional groups needed for the water purification. Also, the reactor operational conditions are adjusted to grow microorganisms in desired aggregate modes. The microbial communities in wastewater treatment are either assembled as flocs in activated sludge or as biofilms on different substrata. Free-floating spherical biofilms, so called granules, combine many of the properties of these two growth modes (Ferrera & Sánchez, 2016, Chen *et al.*, 2017), which has gained increasing interest and implementation in recent years (Zhang *et al.*, 2016). The possibility to remove

carbon, nitrogen and phosphorous simultaneously is one of the most important attributes of granular sludge, which renders highly compact and less energy demanding wastewater treatment (Bengtsson *et al.*, 2018). The granules are compact and spherical suspended biofilms with a diameter of approximately 1-3 mm obtained in defined reactor conditions of high up-flow air velocity, large shear forces, large temporal variation of electron donors and -acceptors causing feast-famine operation and short settling time to select for well settling biomass (Nancharaiah & Kiran Kumar Reddy, 2018, Wilén *et al.*, 2018). Despite well-established methods for granule cultivation, detailed studies on the microbial community ecology in aerobic granules are scarce and the ecological processes underpinning the microbial community assembly during granulation are poorly understood (Sarma *et al.*, 2017).

Laboratory experiments in wastewater treatment are often performed in one reactor due to practical reasons. Hence, conclusions regarding reactor performance and microbial community structure are generally drawn from single reactors operated at different conditions (Weissbrodt *et al.*, 2014, Szabó *et al.*, 2017, Fan *et al.*, 2018, Gonzalez-Martinez *et al.*, 2018). It is unclear how reliable these conclusions are, especially regarding complex processes such as microbial community assembly. Furthermore, laboratory scale wastewater studies that enable controlled environmental conditions and tests of reproducibility are valuable test-benches for studies of microbial ecology (Widder *et al.*, 2016). To better understand he microbial community assembly and reproducibility of granule formation, we have used amplicon sequencing to investigate the succession during the transition from floccular to granular sludge in three replicate reactors seeded with the same inoculum and operated at identical conditions for 35 days. At the same time, the function of the microbial communities was monitored. We hypothesized that similar communities would develop in all three reactors if deterministic factors were dominating, whereas the opposite would support the importance of stochastic factors. We also investigated possible shifts in the importance of deterministic/stochastic factors for community assembly during the operation period.

Material and methods

Experimental set-up and media

Three identical sequencing batch reactors (6 cm diameter and 132 cm height) with a working volume of 3 L were operated for 35 days at room temperature (20-22˚C) in 4-hour cycles (0:05 h filling, 0:55 h anaerobic, 2:53 h aerobic, 0:02 h settling, 0:05 h withdrawal) with a volumetric exchange ratio of 43%. The settling time was gradually decreased to promote granulation and at the same time allow retention of slow growing organisms (Figure S1), and the aerobic phase was adjusted accordingly to permit even 4-hour cycle lengths. Air was provided from the bottom of the reactors with porous diffusers (pore size 1 µm) at a flow rate of 2.5 L min^{-1} and superficial up-flow velocity of 1.5 cm s^{-1}. The reactors were inoculated with aerobic/anoxic activated sludge from a full-scale wastewater treatment plant (Gryaab AB, Gothenburg, Sweden). Synthetic wastewater was pumped in at the bottom of the reactor with a chemical

oxygen demand (COD) to nitrogen ratio of 100:25 (3 kg COD $m^{-3}d^{-1}$ and 0.75 kg NH_4-N $m^{-3}d^{-1}$) and pH of approximately 7 and consisted of 1.6 g L^{-1} NH_4CH_3COO, 12.5 mg L^{-1} $MgSO_4 \cdot 7H_2O$, 15.0 mg L^{-1} $CaCl_2$, 10.0 mg L^{-1} $FeSO_4 \cdot 7H_2O$, 22.5 mg L^{-1} K_2HPO_4, and 1 mL L^{-1} micronutrient solution (0.05 g L^{-1} H_3BO_3, 0.05 g $L^{-1}ZnCl_2$, 0.03 g L^{-1} $CuCl_2$, 0.05 g L^{-1} $MnSO_4.H_2O$, 0.05 g L^{-1} $(NH_4)_6Mo_7O_{24}.4H_2O$, 0.05 g L^{-1} $AlCl_3$, 0.05 g L^{-1} $CoCl_2.6H_2O$, and 0.05 g L^{-1} $NiCl_2$). Biomass samples (100 mL) were withdrawn from the reactors daily for analysis. The replicate reactors were operated identically.

DNA extraction, amplification and sequencing

Samples for microbial community analysis were collected from the reactor three times per week. Total genomic DNA was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals) following the manufacturer's protocol. The V4 region of the 16S rRNA gene was amplified in duplicates, using the forward primer 515'F (5´-GTGBCAGCMGCCGCGGTAA-3´) and the reverse primer 806R (5´-GGACTACHVGGGTWTCTAAT-3´), indexed according to Kozich *et al.* (2013). The PCR reactions were conducted in a 20 µL reaction volume using 17 µL of AccuPrime Pfx SuperMix (Life Technologies), 1 µL of genomic DNA (20 ng template), and 1 µL each of the forward and reverse primers (10 µM). The PCR reactions were carried out in a Biometra T3000 Thermocycler with the following thermal cycling parameters: initial denaturation at 95°C for 5 min, followed by 30 cycles of denaturation (95°C, 20 s), annealing (50°C, 15 s) and elongation (68°C, 60 s), and finished by a 10 min final elongation at 68°C. The duplicate PCR products were pooled, the DNA concentration was normalized, the samples were purified using the Agencourt AMPure system (Beckman Coulter), pooled in equimolar amounts and sequenced using MiSeq reagent kit v3 (Illumina). To increase the confidence in the obtained results, a second MiSeq sequencing run was performed on the same pool. Raw sequencing data were deposited in the NCBI Short Read Archive (SRA), BioProjectID PRJNA472243, accession number SRP148672.

Sequence processing and sequence data analysis

Raw sequence reads were processed following the UNOISE pipeline (Edgar & Flyvbjerg, 2015, Edgar, 2016) with USEARCH v.10 (Edgar, 2010). Briefly, paired-end reads were merged with a maximum number of mismatches of 15 (5% of the read length) and a minimum of 80% of identity in the alignment. Merged reads were filtered (maximum expected error per read of 0.5), dereplicated (default settings) and denoised (minimum abundance of 4), resulting in zero-radius operational taxonomic units (OTUs). OTUs were taxonomically classified with the SINTAX algorithm (Edgar, 2016) based on the MiDAS database v.2.1 (McIlroy *et al.*, 2015) and analysed in R (R Core Team, 2017). The dataset was rarefied, subsampling each sample to 17395 sequences. Basic R functions were used to perform Wilcoxon signed-rank tests and to calculate Pearson correlation coefficients. Non-metric multi-dimensional scaling (NMDS) ordination and heatmaps were created using the R package ampvis (Albertsen *et al.*, 2015). Constrained analysis of proximities (CAP) was performed using vegan R package (Oksanen *et al.*, 2017). The sequences were aligned with the R package DECIPHER (Wright, 2015) and a maximum

likelihood tree was generated with FastTree 2 software (Price *et al.*, 2010) using the GTR+CAT (General Time Reversible with per-site rate CATegories) model of approximation for site rate variation and computation of Gamma20-based likelihood.

α-diversity for each sample was calculated using numbers equivalents (Hill numbers) (Jost, 2006). We refer to these indices as naïve diversity (^{q}ND) because the phylogenetic distance between OTUs are not considered. The parameter q is the diversity order. At a q of 0, all OTUs are considered equally important and, hence, ^{0}ND is the sample richness (i.e. the number of OTUs in a sample). For higher values of q, more weight is put on abundant OTUs. We also calculated diversity numbers which take the sequence dissimilarity into account using the framework described by Chiu & Chao (2014). The metric, here referred to as ^{q}PD (phylogenetic diversity), represents the total effective phylogenetic distance between all OTUs in a sample. For all pairwise comparisons of samples, β-diversity was obtained using the same calculation framework. The β-diversities were converted into dissimilarity indices (^{q}βdis) constrained between 0 (two identical samples) and 1 (two samples with no shared OTUs) (Jost, 2007, Chiu & Chao, 2014). The parameters $^{q}\beta dis_{ND}$ and $^{q}\beta dis_{PD}$ are based on the naïve and the phylogenetic β-diversity values, respectively.

Null model analysis

Null models were used to analyse the taxonomic- and phylogenetic successional turnover between consecutive samples in the reactors. Taxonomic turnover was estimated with Raup-Crick-based measures based on Bray-Curtis dissimilarities (RC_{bray}), calculated with the R package vegan as in Stegen *et al.* (2013). Shortly, a null distribution (999 randomizations) of expected Bray-Curtis dissimilarities was created for each pair of communities, which was compared to the empirically observed Bray-Curtis dissimilarities. The RC_{bray} index is calculated by adding the number of times the null distribution is greater than the empirical Bray-Curtis distribution to half the number of times the null and the empirical distribution are equal. Then this count is normalized to a value between -1 and 1. A negative value means that the two communities share more taxa than expected by chance whereas a positive value means that they share less taxa. Values > |0.95| were considered statistically significant. Values < |0.95| indicates that the taxonomic turnover between the community pair was not different from the null expectation and therefore, influenced by stochastic factors (Zhou & Ning, 2017).

The R package PICANTE (Kembel *et al.*, 2015) was used to calculate the β nearest-taxon index (βNTI), which measures if the phylogenetic turnover is greater or less than the null expectation. For this, the β mean nearest-taxon distance (βMNTD), which measures the mean phylogenetic distance between the most closely related OTUs in two communities, was first calculated based on relative abundance data, as in Stegen *et al.* (2013). In each iteration (999 randomizations) the OTUs were moved randomly across the tips of the regional pool phylogeny and the resulting phylogenetic relationships were used to calculate the $\beta MNTD_{null}$. The βNTI was calculated as the difference in standard deviation units between the observed βMNTD and the mean of

$\beta MNTD_{null}$ in the pairwise sample comparisons. A negative βNTI value means that the samples are more phylogenetically close to each than expected by chance and a positive value means that the two samples are more phylogenetically distant from each other. Previous studies have found that short phylogenetic distances between OTUs, quantified using βMNTD and βNTI, reflect ecological distances (Stegen *et al.*, 2013, Wang *et al.*, 2013, Dini-Andreote *et al.*, 2015). Pairwise comparisons with βNTI > |2| were considered statistically significant. Values < |2| were not significantly different from the null expectation, which indicate that stochastic factors influenced the phylogenetic turnover (Zhou & Ning, 2017).

Analytical methods

Concentrations in the effluent of total organic carbon (TOC) and total nitrogen (TN) were measured with a Shimadzu TOC analyser. Total suspended solids and volatile suspended solids in the in the reactor (MLSS, MLVSS) and in the effluent (effTSS, effVSS) were measured according to standard methods. Microscopy was performed using an Olympus BX60 light microscope (Olympus Sverige AB, Solna, Sweden) with particle size assessment by measuring the diameter of 10 random flocs/granules using the CellSens (Olympus) software.

Results

Granulation of the sludge and reactor performance

The sludge started to granulate around day 7, when small granules emerged (Figure S2) and the mean particle sizes of the reactor biomass increased similarly in the three reactors (Figure S3). At day 25, the granulated biomass dominated in the reactors and at the end of the experiment (day 35), the granules had an average diameter of 1.5 mm. The three reactors also followed similar patterns over time in suspended solids concentrations in the reactor and in the effluent (Figure 1A-B), although R1 had significantly lower reactor concentrations (MLSS and MLVSS) than R2 and R3 ($p<0.05$) (Table S1). Removal of TOC and TN was similar among the reactors (Figure 1, C-D). TOC removal was above 90% since day 4, with an average of 95% for R1 and R3 and 96% for R2. The average TN removal was 54%, 56% and 59% in R1, R2 and R3 respectively and increased to above 70% from day 28 for R1 and R3, and day 32 for R2.

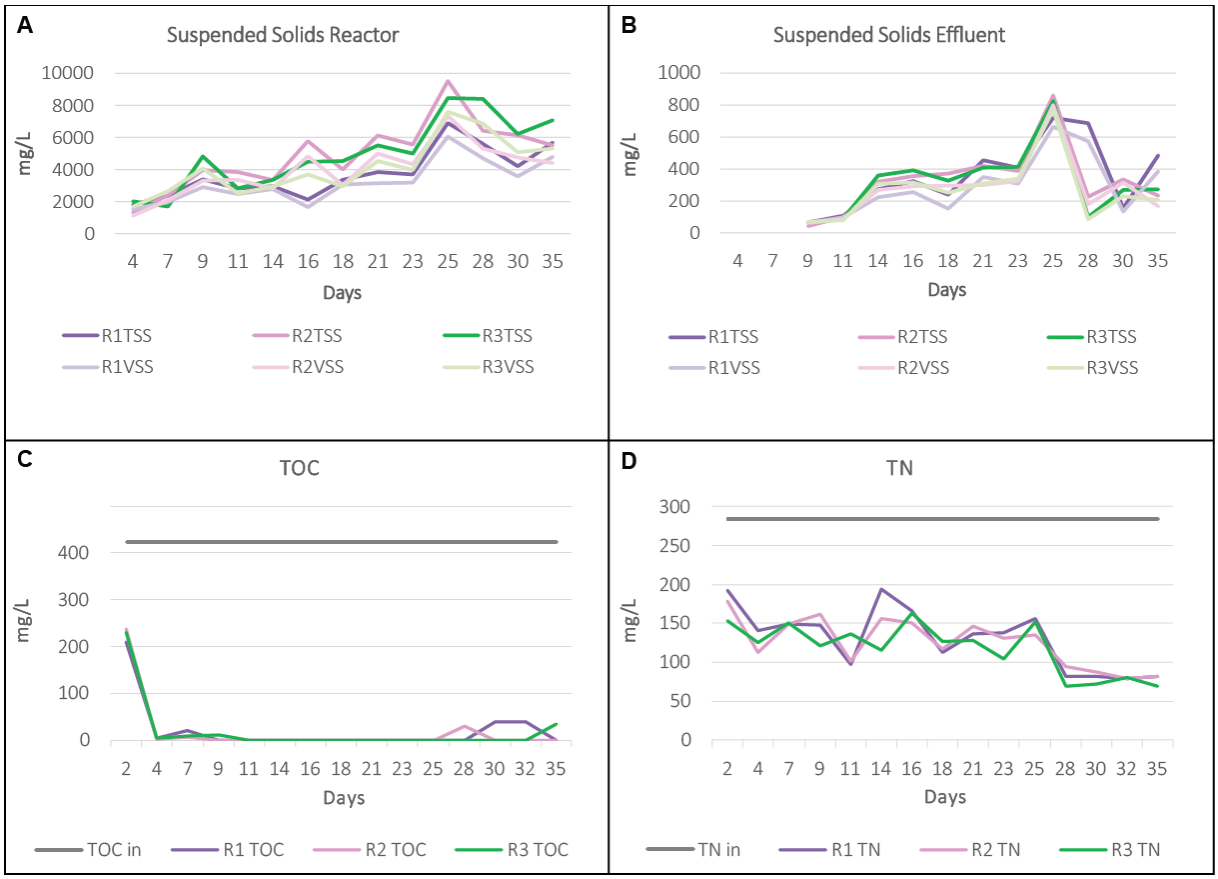

Figure 1. Sludge properties and reactor performances over time for R1 (purple), R2 (pink) and R3 (green). A, reactor suspended solids concentrations; B, effluent suspended solids concentrations; C, TOC concentrations in the influent and effluent; D, TN concentrations in the influent and effluent.

Temporal variation of abundant taxa

The microbial community composition was highly dynamic during the transition from floccular to granular sludge. The seed sludge, mainly composed of *Betaproteobacteria* (38% of total reads), *Alphaproteobacteria* (14%), *Sphingobacteria* (11%) and *Clostridia* (7%) at class level, changed rapidly after inoculation to a community dominated by *Gammaproteobacteria* (70%) with *Acinetobacter* as the dominant genus during the first 7 days (Figure 2, S4). Thereafter, *Betaproteobacteria* dominated in the three reactors, with high abundances of the genera *Thauera*, *Acidovorax*, *Simplicispira*, *Comamonas*, *Curvibacter* and *Alicycliphilus* (Figure 2). From days 21-25, *Actinobacteria* dominated until the end of the experiment, with *Corynebacterium* sp. as only representative of this class, having more than 30% of the total reads in the three reactors. Many taxa that were rare or undetected in the seed sludge became abundant during the granulation, e.g. *Corynebacterium* sp., *Brevundimonas* sp., *Pedobacter* sp., *Fluviicola* sp. and *Bacteriovorax* sp., while other taxa that were common in the seed sludge decreased drastically in relative abundance, e.g. *Rhodoferax* sp., *Rhodobacter* sp., *Pseudorhodobacter* sp. and *Hydrogenophaga* sp.

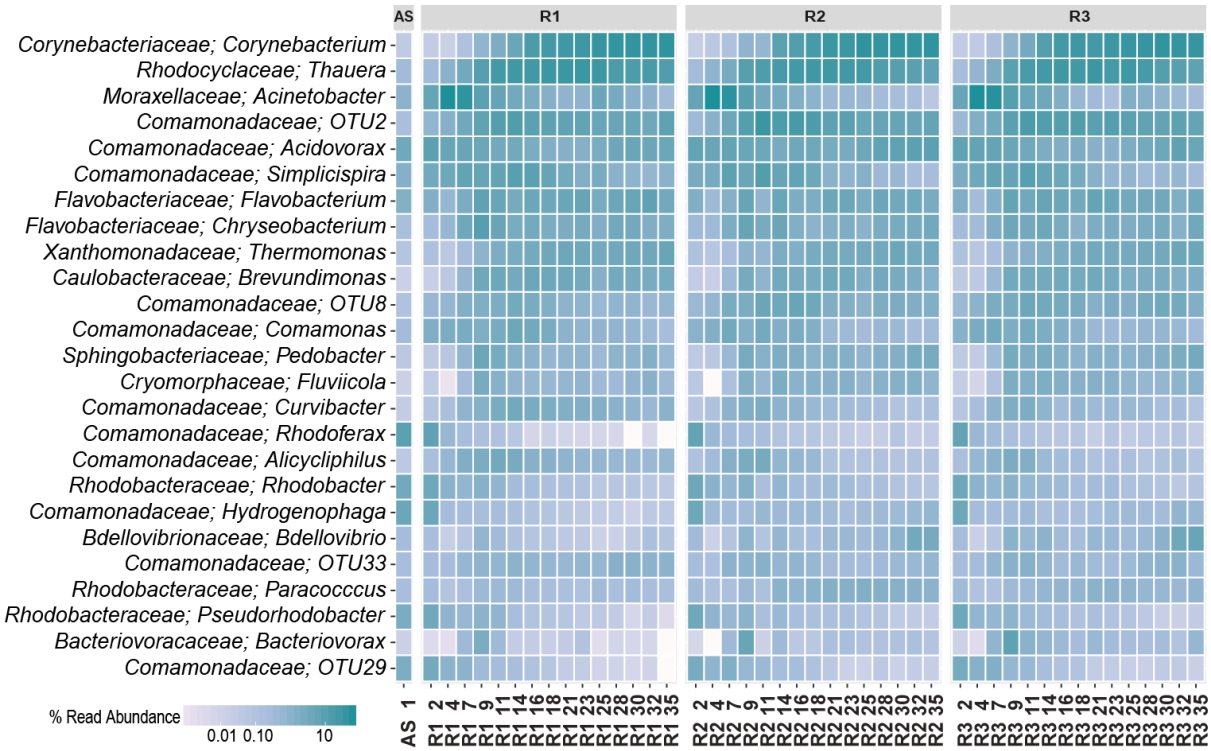

Figure 2. Temporal variation of the 25 most abundant genera as percentages of reads in the replicate reactors (R1, R2 and R3) and the seed sludge (AS), from days 1 to 35.

Microbial diversity and dynamics during granulation

The three reactors followed the same trends in α-diversity (Figure 3A-C). After inoculation, there was a drop in diversity, followed by a peak around day 9 and then a continuous decrease. R1 differed somewhat from R2 and R3 during the dynamic period after the inoculation and during days 21-35 when ^{0}ND was consistently lower in R1, while ^{1}ND and ^{2}ND were similar in the three reactors, showing that the rare OTUs were fewer in R1. The observed trends for naïve diversity were very similar to the trends for the phylogenetic diversity (Figure S5A-C).

Succession, measured as β-diversity between successive sample points, showed that after inoculation there were large temporal dynamics in each reactor with dissimilarities of around 0.5 for ^{0}βdisND and even higher peak dissimilarities (0.8) for ^{2}βdisND, which gives large weight to abundant OTUs (Figure 3D-F). Thereafter, the dissimilarities between successive samples decreased considerably to values of about 0.3 for ^{0}βdisND and 0.05 for ^{1}βdisND and ^{2}βdisND.

For β-diversity between the reactors (Figure 3G-I), ^{0}βdisND, was around 0.3-0.4, showing that 60-70% of the OTUs were shared among the reactors at any given time. ^{1}βdisND was around 0.1 and ^{2}βdisND around 0.05, i.e. the reactors were more similar to each other when considering the more abundant OTUs. R2-R3 were more similar than R1-R2 or R1-R3 for diversity order 0 and 1, but not for diversity order 2. Also for β-diversity between successive

sampling points or between reactors, the naïve diversity patterns were highly similar to the ones for phylogenetic diversity (Figure S5).

NMDS shows initial, rapid changes after inoculation and that R1 and R2 deviated from the other reactors at day 7 and day 11 (Figure 4). From day 14, R2 and R3 converged and evolved similarly after that, whereas R1 diverged. CAP analysis based on Bray-Curtis dissimilarities confirmed R1 to be different to R2 and R3 (Figure S6). Pearson correlation based on CAP showed that 144 out of the 2540 OTUs were significantly different ($p<0.05$) in relative abundance between reactors (Supplementary material). 58 OTUs were more abundant in R1 than R2-R3 and 82 OTUs more abundant in R2-R3 than R1. These OTUs, mainly of low relative abundance added up to 8% of the total reads.

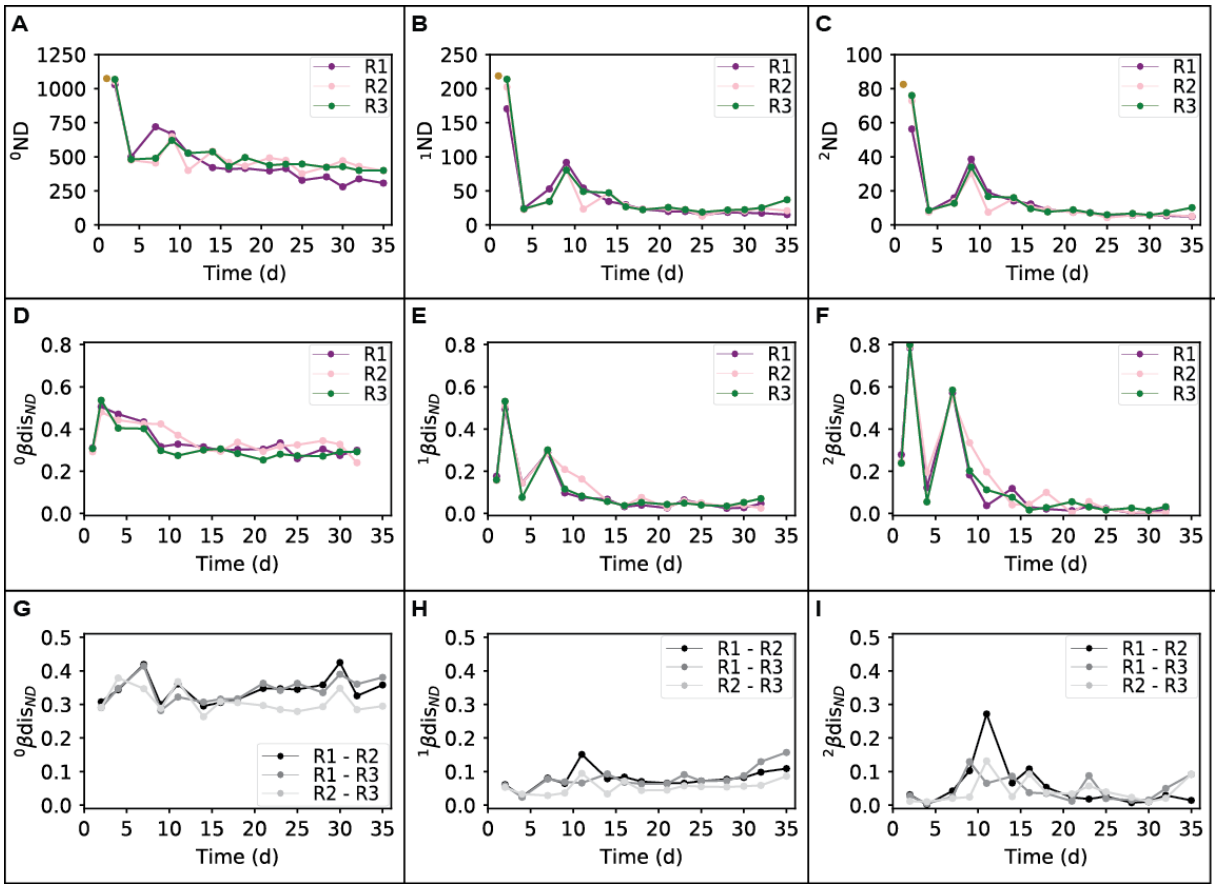

Figure 3. Dynamics of the naïve diversity (taxonomic distance between OTUs) for the replicate reactors. A, B and C: naïve α-diversity; C, D and F: naïve β-diversity over time between two successive sample points; G, H and I: naïve β-diversity over time between the replicate reactors.

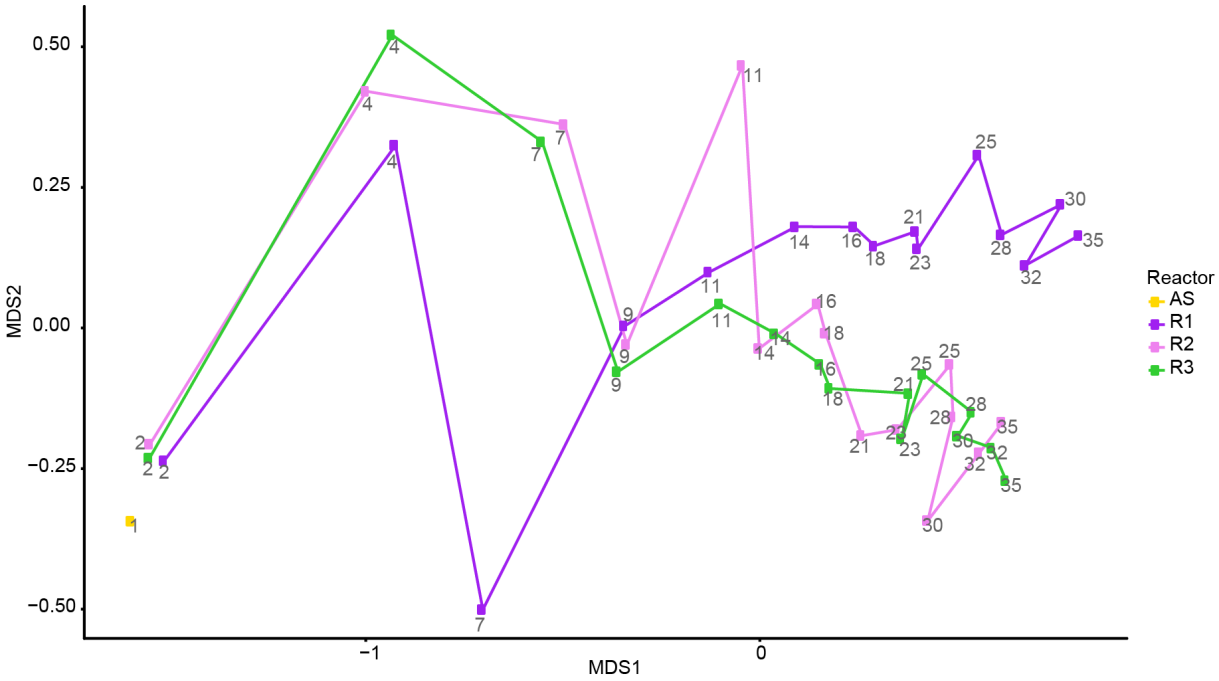

Figure 4. NMDS ordination based on Bray-Curtis dissimilarities after square root transformation. Each line represents the community succession of each reactor, R1 (purple), R2 (pink), R3 (green), and the seed sludge (AS, day 1, yellow). Numbers refer to days after startup. Stress = 0.047.

Null model analysis

We wished to determine the factors behind the shifts in β-diversity over time in each reactor. However, α-, β- and γ-diversity are interconnected so that e.g. a difference in richness between samples may in itself cause a difference in β-diversity between the samples (Chase *et al.*, 2011, Zhou *et al.*, 2013). Null model analysis was therefore used to disentangle turnover due to succession from changes in α- and γ-diversity. We used two different types of null models for this, RC_{bray}, based on taxonomic turnover, and βNTI, based on phylogenetic turnover. Turnover between sequential samples from the same reactor would show successional phases. Thus, RC_{bray} and βNTI values lower than the null expectation, would indicate "steady-state" microbial community composition. In R1, RC_{bray} values were close to 1 during the first 9 days (Figure 5A), indicating higher taxonomic turnover than the null expectation, and thereafter turnover remained lower than the null expectation (RC_{bray} < -0.95). R2 an R3 showed a somewhat different trend with turnover closer to the null expectation (RC_{bray} between 0.95 and -0.95) during the first 11 days and after this period, the taxonomic turnover between successive sample points remained lower than expected by chance (RC_{bray} < -0.95). Similar patterns were observed using βNTI, with higher phylogenetic turnover for R1 than for R2 and R3 until day 14, followed by lower turnover than expected by chance in all reactors (Figure S7). RC_{bray} was also used for the fraction of OTUs having significantly different relative abundance between the reactors, as determined by correlations to the CAP axes (Figure S6 and Supplementary material). For this fraction of the community, the taxonomic turnover between succesive

sample points in R2 and R3 gradually increased from day 9 to values not significantly different to the null expectation whereas in R1, the taxonomic turnover remained lower than the null expectation throughout the reactor operation (Figure 5B).

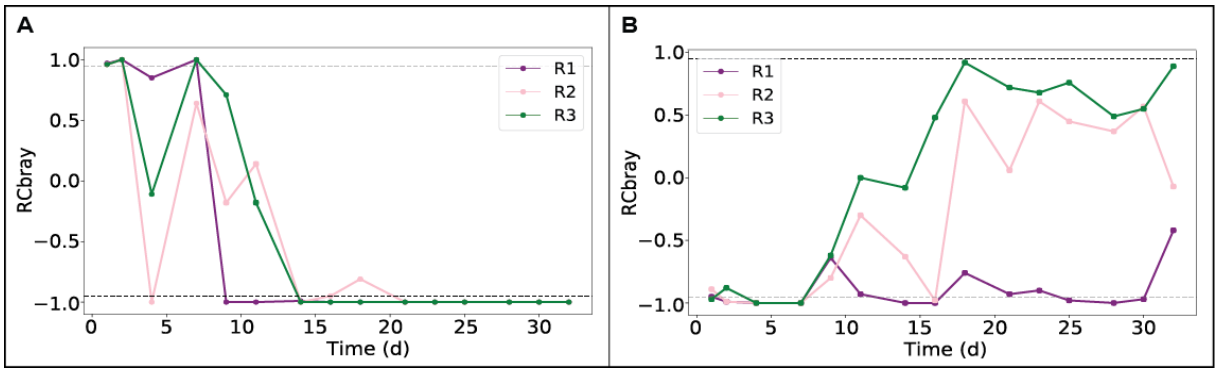

Figure 5. Taxonomic turnover (RC_{bray}) in community composition between two succesive sample points of the whole community (A) and of OTUs with a significant p-value ($p<0.05$) in the Pearson correlation for the CAP2 axis (B) of the replicate reactors.

Sequencing reproducibility

To ensure that the observed differences between the reactors, mainly caused by OTUs of low abundance, was not caused by limited precision of sequencing, the samples were sequenced twice. Highly similar results were obtained from the two sequencing runs with regard to relative read abundance of major OTUs, NMDS ordination and CAP (Figures S8-S10).

Discussion

In this study, we investigated how communities of microorganisms assembled over time as granules formed in three replicate reactors for wastewater treatment. Granule growth, successional patterns of taxa, microbial diversity and reproducibility are discussed together with an assessment of the mechanisms that influence the community assembly.

Succession of taxa during the granulation process

During the early transformation from flocs to granule-like particles, *Acinetobacter* sp. became increasingly abundant in all reactors (Figure 2). *Acinetobacter* sp. has been reported as primary colonizers and bridging bacteria in aggregate/biofilm development (Katharios-Lanwermeyer *et al.*, 2014, Liébana *et al.*, 2016) and have previously been observed during initial stages of granulation (Weissbrodt *et al.*, 2012). *Acinetobacter* sp. is a hydrophobic bacterium and produces EPS, which are decisive properties for early stage aggregate formation (Ding *et al.*, 2015) and likely explains their frequent occurrence during the initial phase.

As the granules started to increase in size, after one week of operation (Figure S3), the microbial communities shifted and bacteria within *Comamonadaceae*, *Rhodocyclaceae*, *Flavobacteriaceae*, *Xanthomonadaceae* and *Caulobacteraceae* became more abundant. Many

of these bacteria are associated with granulation (Weissbrodt *et al.*, 2012, Szabó *et al.*, 2017). *Thauera* sp., which was particularly abundant in this period (Figure 2), is an important EPS producer commonly found in acetate-fed aerobic granules and is also a well-known denitrifier, as are many of the other common microorganisms in this period (Szabó *et al.*, 2017, Xia *et al.*, 2018). For instance, bacteria within the *Comamonadaceae* family are considered important for acetate-fed denitrification in aerobic granules (Lu *et al.*, 2014). Despite this, the total nitrogen removal was incomplete, around 50% until day 28. Thereafter nitrogen removal increased to 70% until the end of the study period, day 35. In this last period, *Corynebacterium* sp. was highly abundant in the reactors (Figure 2). *Corynebacteria* can remove nitrogen both by assimilatory reduction and denitrification (Seenivasagan *et al.*, 2014).

None of the common members of the microbial community in the granules were highly abundant, or even detected in the seed sludge. This is commonly observed in reactor studies (Vanwonterghem *et al.*, 2014, Szabó *et al.*, 2017, Fan *et al.*, 2018) and experimental evidence suggests that during succession, rare taxa can be promoted to the dominant fraction of the microbial community following a perturbation or a change of environment (Pagaling *et al.*, 2013, Shade *et al.*, 2014, Coveley *et al.*, 2015).

Microbial community assembly during granulation

The microbial communities in the replicate reactors followed similar succession patterns but with a clear divergence of R1 from R2 and R3 (Figure 4). Since the inoculum, feed composition, and reactor operation were the same among the reactors, this suggests that stochastic factors had a role in the community assembly. Let us first consider the two extreme cases of either stochastic or deterministic factors acting alone. With only stochastic factors, i.e. drift, neutral theory predicts that the microbial communities in the three reactors would eventually diverge and the dissimilarity index between reactors would increase and reach unity. On the other hand, if deterministic factors, i.e. selection, were acting alone, the reactors will follow the same trajectory and the dissimilarity index between reactors would be zero. The ecological processes responsible for microbial community assembly can be classified as selection, dispersal, drift, and diversification (Vellend, 2010). Our experimental design excluded dispersal as a major factor because synthetic feed was used. Diversification can most likely also be neglected because of the short time frame of the study (35 days). This leaves selection as the major deterministic process and drift as the major stochastic process affecting community assembly.

During the early start-up, the microbial community shifted rapidly from the inoculum in a similar manner in all reactors (Figure 4) and the phylogenetic successional turnover was relatively high, especially in R1 (Figure S7). The settling time for the biomass was long (Figure S1), thus the microorganisms were not washed out due to their settling properties, but according to their ability to grow during these reactor conditions (Szabó *et al.*, 2017). Therefore, the changes in the microbial community structure, resulting in a harsh decrease in diversity (Figure 3), were to a large extent caused by the switch from municipal wastewater to a synthetic

feed with acetate as carbon source. A switch from a complex substrate to a simple and easily biodegradable substrate would favour specific taxa, and consequently, the biodiversity, both taxonomic and phylogenetic, would decrease (Weissbrodt *et al.*, 2012, Szabó *et al.*, 2017). Indeed, the drop in diversity was particularly evident for highly abundant OTUs (^{2}ND and ^{2}PD, Figure 3, S5) and the dissimilarity between the consecutive days 2 and 4 was also higher for the more abundant OTUs (^{2}βdisND and ^{2}βdisPD, Figure 3). In addition, the lack of immigration of new taxa (dispersal) to the reactors due the synthetic feed probably contributed to the decrease in α-diversity (Albright & Martiny, 2017).

The settling time was thereafter drastically reduced, from 30 min to 5 min between days 5 and 11, to promote granulation (Figure S1) and the first granules were observed at day 7. This led to increased α-diversity in the reactors as new, previously undetected, OTUs grew in relative abundance (Figure 3), presumably due to the formation of new niches in the emerging granules. The communities changed rapidly over time, in particular around day 7 (Figure 3), and it was clear that the community assembly took different paths in the three reactors in this phase (Figure 4). The null-models indicated that stochastic factors, here mainly drift, were important for the succession during this dynamic phase (Figure 5, S7). This could be due to the random distribution of microorganisms between aggregates of different sizes of which only the larger and denser ones were retained by sedimentation (Szabó *et al.*, 2017).

From day 11, the decrease in settling time was less dramatic (Figure S1) and the granules increased gradually in size (Figure S3). In this phase, the microbial communities changed more slowly with time (Figure 3, S5) and turnover was lower than predicted by chance by the null-models (Figure 5, S7), suggesting that deterministic forces, i.e. selection, were important. The microbial succession in R2 and R3 followed the same trajectory until the end of the experiment, but R1 diverged (Figure 4). The divergence of R1 was not due to the most abundant OTUs, the community was dominated by the same genera in the three replicate reactors (Figure 2). However, α-diversity (^{0}ND) was lower in R1 than R2 and R3 and β-diversity between reactors (Figure 3) showed that the divergence of R1 was mainly caused by rare and intermediate OTUs with relative abundance below 1% (See OTU table of OTUs with a significant p-value ($p<0.05$) in the Pearson correlation for the CAP2 axis in supplementary material). A subset of the OTUs were significantly different among the reactors, as assessed by correlation to the CAP axes (Figure S6). For these OTUs, successional turnover was lower than by chance in R1, while for R2 and R3, the null-model predicted stochasticity (drift) to have larger influence (Figure 5B). Within the same microbial community, selection and drift can affect different subpopulations differently (Zhou & Ning, 2017). The less abundant OTUs were likely more susceptible to drift, since many events could result in extinction (Nemergut *et al.*, 2013). Interestingly, predatory bacteria belonging to the genus *Bdellovibrio*, *Bacteriovorax* and *Peredibacter* were detected in the three reactors at as high as 18% maximum relative abundance. *Bdellovibrio* and like organisms (BALOs) prey on gram negative bacteria, with some BALOs being generalist and some specialists in their prey preferences (Johnke *et al.*, 2014). BALOs have been found in granules (Li *et al.*, 2014, Wan *et al.*, 2014, Weissbrodt *et al.*, 2014) and have recently been shown to

affect the granular microbial structure (Feng *et al.*, 2017). R2 and R3 harboured considerably more BALOs than R1, especially in the later stages of the study (Figure 2 and Figure S8). Also, out of a total of 60 OTUs belonging to BALOs, 11 were significantly lower in abundance in R1 compared to R2 and R3 according to CAP (Supplementary material). It could be hypothesized that BALOs caused a higher turnover in R2 and R3, contributing to the divergence of these reactors from R1.

We tentatively interpret the difference between the reactors in terms of underlying mechanisms: During the early phase the change of feed selected for certain taxa, which caused the biodiversity to decrease (Figure 3). The successional turnover was high, but the microbial communities in replicate reactors were similar (Figure 2 and 4), therefore the acclimation to a new feed seems to be mostly deterministic. A second transitional phase occurred when the settling time was decreased and granules began to emerge. This likely created new niches, which allowed low-abundant taxa, previously below the detection threshold, to increase in abundance resulting in temporarily increased α-diversity (Figure 3). During this phase, stochasticity seemed to cause the microbial community in reactor 1 to take a different path (Figure 4). However, selection eventually led to the same dominating taxa in the parallel reactors (Figure 2). The successional turnover was low during this last period, presumably because selection had led to the development of a microbial community adapted to the specific reactor conditions, but some drift caused reactors 2 and 3 to slowly diverge from reactor 1.

Ecological processes and wastewater treatment

Biological processes for wastewater treatment are often tested in laboratory- or pilot-scale bioreactors. Bioreactors are designed as if they were solely affected by deterministic processes and the treatment process relies on this assumption. But the literature is inconsistent regarding the reproducibility of wastewater biological reactors. Vanwonterghem *et al.* (2014) observed that the microbial community structure and dynamics in three replicate anaerobic digesters were highly reproducible, influenced by deterministic processes. Ayarza *et al.* (2010) reported the microbial community dynamics to follow similar patterns in four replicate sequencing batch reactors treating synthetic wastewater due to deterministic processes, although, drift caused the reactors to display a differences. Zhou *et al.* (2013) demonstrated divergent microbial communities in 14 replicate microbial electrolysis cell reactors treating wastewater due to a dominant role of stochastic processes. Despite highly dynamic microbial communities , stable performance of parallel bioreactors treating wastewater are often found for such general functions as aerobic/anoxic oxidation of organic matter in activated sludge (Kaewpipat & Grady, 2002, Ayarza & Erijman, 2011, Bagchi *et al.*, 2015, Goux *et al.*, 2015). On the other hand, for complex processes that require multispecies cooperation, such as anaerobic digestion or specialized processes such as microbial electrolysis, stochastic variations in community composition has resulted in altered microbial functions (Zhou *et al.*, 2013, Goux *et al.*, 2015). In the present study the differences in the microbial community dynamics were due to differences in the abundance of a small fraction of the total OTUs, which did not have major

impact on the reactor functions since the replicate reactors behaved similarly (Figure 1). The same dynamics in dominant OTUs and a higher variability of the less abundant fraction of the community has also been observed in full-scale activated sludge reactors (Griffin & Wells, 2016). From the engineering perspective our results are important since they not only confirm the possibility to select for certain dominating taxa by manipulating the reactors conditions, even for a process as common as acetate oxidation, but also provide information about fundamental ecological processes involved in sludge granulation. Moreover, the results presented here suggest that it is possible to draw conclusions about the performance and microbial community dynamics (at least for abundant community members) in experiments with single aerobic granular sludge reactors.

In summary, our study shows that three parallel granular sludge reactors operated identically and inoculated with the same seed sludge behaved similarly, both in terms of treatment performance and microbial community assembly. Deterministic and stochastic processes exerted a dynamic influence during the different successional stages. Deterministic factors shaped the community composition of abundant members whereas stochastic factors mostly affected rare members of the communities. Drift caused the microbial community in one of the reactors to diverge from the other two. Predation seems to have an influence in the divergence of the reactors. Despite these differences in the microbial community, the reactors functions were stable and similar between reactors. More studies dealing with fundamental ecological processes involved in granulation are needed. For instance, it is also important to know how perturbations affect the microbial community and the resilience of the system. This knowledge will help to better understand the technology which ultimately will improve its application and management.

Acknowledgements

This work was supported by FORMAS (Contract no. 245-2013-627), the Swedish Research Council for Environment, Agriculture Sciences and Spatial Planning. The authors acknowledge the Centre for Cellular Imaging and the Genomics core facility at the University of Gothenburg for support and use of their equipment.

REFERENCES

Albertsen M, Karst SM, Ziegler AS, Kirkegaard RH & Nielsen PH (2015) Back to Basics–The Influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. *PloS one* **10**: e0132783.

Albright MBN & Martiny JBH (2017) Dispersal alters bacterial diversity and composition in a natural community. *The Isme Journal* **12**: 296.

Anderson MJ, Crist TO, Chase JM*, et al.* (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters* **14**: 19-28.

Ayarza JM & Erijman L (2011) Balance of neutral and deterministic components in the dynamics of activated sludge floc assembly. *Microbial Ecology* **61**: 486-495.

Ayarza JM, Guerrero LD & Erijman L (2010) Nonrandom Assembly of Bacterial Populations in Activated Sludge Flocs. *Microbial Ecology* **59**: 436-444.

Bagchi S, Tellez B, Rao H, Lamendella R & Saikaly P (2015) Diversity and dynamics of dominant and rare bacterial taxa in replicate sequencing batch reactors operated under different solids retention time. *Applied Microbiology and Biotechnology* **99**: 2361-2370.

Bengtsson S, de Blois M, Wilén B-M & Gustavsson D (2018) A comparison of aerobic granular sludge with conventional and compact biological treatment technologies. *Environmental Technology* 1-10.

Chase JM, Kraft NJB, Smith KG, Vellend M & Inouye BD (2011) Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* **2**: 1-11.

Chen Y, Lan S, Wang L, Dong S, Zhou H, Tan Z & Li X (2017) A review: Driving factors and regulation strategies of microbial community structure and dynamics in wastewater treatment systems. *Chemosphere* **174**: 173-182.

Chiu C-H & Chao A (2014) Distance-Based Functional Diversity Measures and Their Decomposition: A Framework Based on Hill Numbers. *PLOS ONE* **9**: e100014.

Coveley S, Elshahed MS & Youssef NH (2015) Response of the rare biosphere to environmental stressors in a highly diverse ecosystem (Zodletone spring, OK, USA). *PeerJ* **3**: e1182.

Ding Z, Bourven I, Guibaud G, van Hullebusch ED, Panico A, Pirozzi F & Esposito G (2015) Role of extracellular polymeric substances (EPS) production in bioaggregation: application to wastewater treatment. *Applied Microbiology and Biotechnology* **99**: 9883-9905.

Dini-Andreote F, Stegen JC, van Elsas JD & Salles JF (2015) Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences* **112**: E1326-E1332.

Edgar R (2016) SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv*.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics (Oxford, England)* **26**: 2460-2461.

Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv*.

Edgar RC & Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics (Oxford, England)* **31**: 3476-3482.

El-Chakhtoura J, El-Fadel M, Rao HA, Li D, Ghanimeh S & Saikaly PE (2014) Electricity generation and microbial community structure of air-cathode microbial fuel cells powered with the organic fraction of municipal solid waste and inoculated with different seeds. *Biomass and Bioenergy* **67**: 24-31.

Falk MW, Song KG, Matiasek MG & Wuertz S (2009) Microbial community dynamics in replicate membrane bioreactors - Natural reproducible fluctuations. *Water Research* **43**: 842-852.

Fan X-Y, Gao J-F, Pan K-L, Li D-C, Zhang L-F & Wang S-J (2018) Shifts in bacterial community composition and abundance of nitrifiers during aerobic granulation in two nitrifying sequencing batch reactors. *Bioresource Technology* **251**: 99-107.

Feng S, Tan CH, Constancias F, Kohli GS, Cohen Y & Rice SA (2017) Predation by Bdellovibrio bacteriovorus significantly reduces viability and alters the microbial community composition of activated sludge flocs and granules. *FEMS Microbiology Ecology* **93**: fix020-fix020.

Ferrera I & Sánchez O (2016) Insights into microbial diversity in wastewater treatment systems: How far have we come? *Biotechnology Advances* **34**: 790-802.

Gonzalez-Martinez A, Muñoz-Palazon B, Maza-Márquez P, Rodriguez-Sanchez A, Gonzalez-Lopez J & Vahala R (2018) Performance and microbial community structure of a polar Arctic Circle aerobic granular sludge system operating at low temperature. *Bioresource Technology* **256**: 22-29.

Goux X, Calusinska M, Lemaigre S, Marynowska M, Klocke M, Udelhoven T, Benizri E & Delfosse P (2015) Microbial community dynamics in replicate anaerobic digesters exposed sequentially to increasing organic loading rate, acidosis, and process recovery. *Biotechnology for Biofuels* **8**: 122.

Griffin JS & Wells GF (2016) Regional synchrony in full-scale activated sludge bioreactors due to deterministic microbial community assembly. *The Isme Journal* **11**: 500.

Han S, Liu Y, Zhang S & Luo G (2016) Reactor performances and microbial communities of biogas reactors: effects of inoculum sources. *Applied Microbiology and Biotechnology* **100**: 987-995.

Johnke J, Cohen Y, de Leeuw M, Kushmaro A, Jurkevitch E & Chatzinotas A (2014) Multiple micro-predators controlling bacterial communities in the environment. *Current Opinion in Biotechnology* **27**: 185-190.

Jost L (2006) Entropy and diversity. *Oikos* **113**: 363-375.

Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology* **88**: 2427-2439.

Kaewpipat K & Grady CPL (2002) Microbial population dynamics in laboratory-scale activated sludge reactors. *Water Science and Technology* **46**: 19-27.

Katharios-Lanwermeyer S, Xi C, Jakubovics NS & Rickard AH (2014) Mini-review: Microbial coaggregation: ubiquity and implications for biofilm development. *Biofouling* **30**: 1235-1251.

Kembel S, Ackerly D, Blomberg S, Cornwell W, Cowan P, Helmus M, Morlon H & Webb C (2015) Picante: R tools for integrating phylogenies and ecology. R package version 1.6-2. p.^pp.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK & Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and environmental microbiology* **79**: 5112-5120.

Li Y, Zou J, Zhang L & Sun J (2014) Aerobic granular sludge for simultaneous accumulation of mineral phosphorus and removal of nitrogen via nitrite in wastewater. *Bioresource Technology* **154**: 178-184.

Liébana R, Arregui L, Santos A, Murciano A, Marquina D, Serrano S & Stams A (2016) Unravelling the interactions among microbial populations found in activated sludge during biofilm formation. *FEMS Microbiology Ecology* **92**: fiw134-fiw134.

Lu H, Chandran K & Stensel D (2014) Microbial ecology of denitrification in biological wastewater treatment. *Water Res* **64**.

Lucas R, Kuchenbuch A, Fetzer I, Harms H & Kleinsteuber S (2015) Long-term monitoring reveals stable and remarkably similar microbial communities in parallel full-scale biogas reactors digesting energy crops. *FEMS Microbiology Ecology* **91**: fiv004-fiv004.

Luo G, De Francisci D, Kougias PG, Laura T, Zhu X & Angelidaki I (2015) New steady-state microbial community compositions and process performances in biogas reactors induced by temperature disturbances. *Biotechnology for Biofuels* **8**: 3.

McIlroy SJ, Saunders AM, Albertsen M, Nierychlo M, McIlroy B, Hansen AA, Karst SM, Nielsen JL & Nielsen PH (2015) MiDAS: the field guide to the microbes of activated sludge. *Database* **2015**.

McMahon KD, Martin HG & Hugenholtz P (2007) Integrating ecology into biotechnology. *Current Opinion in Biotechnology* **18**: 287-292.

Nancharaiah YV & Kiran Kumar Reddy G (2018) Aerobic granular sludge technology: Mechanisms of granulation and biotechnological applications. *Bioresource Technology* **247**: 1128-1143.

Nemergut DR, Schmidt SK, Fukami T*, et al.* (2013) Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews* **77**: 342-356.

Oksanen F, Blanchet G, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin P, O'Hara R, Simpson G & Solymos P (2017) vegan: Community Ecology Package. R package version 2.4-4. http s. *CRAN R-pro j ect org/pack age= vega n*.

Pagaling E, Strathdee F, Spears BM, Cates ME, Allen RJ & Free A (2013) Community history affects the predictability of microbial ecosystem development. *The Isme Journal* **8**: 19.

Price MN, Dehal PS & Arkin AP (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE* **5**: e9490.

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016. p.^pp.

Sarma SJ, Tay JH & Chu A (2017) Finding Knowledge Gaps in Aerobic Granulation Technology. *Trends in Biotechnology* **35**: 66-78.

Seenivasagan R, Rajakumar S, Kasimani R & Ayyasamy PM (2014) Screening of assimilatory and dissimilatory denitrifying microbes isolated from nitrate-contaminated water and soil. *Preparative Biochemistry and Biotechnology* **44**: 586-597.

Shade A (2016) Diversity is the question, not the answer. *The Isme Journal* **11**: 1.

Shade A, Jones SE, Caporaso JG, Handelsman J, Knight R, Fierer N & Gilbert JA (2014) Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. *mBio* **5**.

Solli L, Håvelsrud OE, Horn SJ & Rike AG (2014) A metagenomic study of the microbial communities in four parallel biogas reactors. *Biotechnology for Biofuels* **7**: 146.

Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, Rockhold ML & Konopka A (2013) Quantifying community assembly processes and identifying features that impose them. *The Isme Journal* **7**: 2069.

Szabó E, Liébana R, Hermansson M, Modin O, Persson F & Wilén B-M (2017) Microbial Population Dynamics and Ecosystem Functions of Anoxic/Aerobic Granular Sludge in Sequencing Batch Reactors Operated at Different Organic Loading Rates. *Frontiers in Microbiology* **8**: 770.

Szabó E, Liébana R, Hermansson M, Modin O, Persson F & Wilén B-M (2017) Microbial population dynamics and ecosystem functions of anoxic/aerobic granular sludge in sequencing batch reactors operated at different organic loading rates. *Front Microbiol* **8**.

Szabó E, Liébana R, Hermansson M, Modin O, Persson F & Wilén B-M (2017) Comparison of the bacterial community composition in the granular and the suspended phase of sequencing batch reactors. *AMB Express* **7**: 168.

Wan C, Yang X, Lee D-J, Liu X & Sun S (2014) Partial nitrification using aerobic granule continuous-flow reactor: Operations and microbial community. *Journal of the Taiwan Institute of Chemical Engineers* **45**: 2681-2687.

Wang J, Shen J, Wu Y, Tu C, Soininen J, Stegen JC, He J, Liu X, Zhang L & Zhang E (2013) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J* **7**: 1310-1321.

Vanwonterghem I, Jensen PD, Dennis PG, Hugenholtz P, Rabaey K & Tyson GW (2014) Deterministic processes guide long-term synchronised population dynamics in replicate anaerobic digesters. *ISME J* **8**: 2015-2028.

Weissbrodt DG, Shani N & Holliger C (2014) Linking bacterial population dynamics and nutrient removal in the granular sludge biofilm ecosystem engineered for wastewater treatment. *FEMS Microbiology Ecology* **88**: 579-595.

Weissbrodt DG, Lochmatter S, Ebrahimi S, Rossi P, Maillard J & Holliger C (2012) Bacterial selection during the formation of early-stage aerobic granules in wastewater treatment systems operated under wash-out dynamics. *Front Microbiol* **3**: 332.

Vellend M (2010) Conceptual synthesis in community ecology. *The Quarterly review of biology* **85**: 183-206.

Widder S, Allen RJ, Pfeiffer T*, et al.* (2016) Challenges in microbial ecology: building predictive understanding of community function and dynamics. *The Isme Journal* **10**: 2557.

Wilén B-M, Liébana R, Persson F, Modin O & Hermansson M (2018) The mechanisms of granulation of activated sludge in wastewater treatment, its optimization, and impact on effluent quality. *Applied Microbiology and Biotechnology* 1-16.

Wittebolle L, Van Vooren N, Verstraete W & Boon N (2009) High reproducibility of ammonia-oxidizing bacterial communities in parallel sequential batch reactors. *Journal of Applied Microbiology* **107**: 385-394.

Wright ES (2015) DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics* **16**: 322.

Xia J, Ye L, Ren H & Zhang X-X (2018) Microbial community structure and function in aerobic granular sludge. *Applied Microbiology and Biotechnology* **102**: 3967-3979.

Xia Y, Wen X, Zhang B & Yang Y (2018) Diversity and assembly patterns of activated sludge microbial communities: A review. *Biotechnology Advances* **36**: 1038-1047.

Zhang Q, Hu J & Lee D-J (2016) Aerobic granular processes: Current research trends. *Bioresource Technology* **210**: 74-80.

Zhou J & Ning D (2017) Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiology and Molecular Biology Reviews* **81**.

Zhou J, Liu W, Deng Y, Jiang Y-H, Xue K, He Z, Van Nostrand JD, Wu L, Yang Y & Wang A (2013) Stochastic Assembly Leads to Alternative Communities with Distinct Functions in a Bioreactor Microbial Community. *mBio* **4**.