Press xenobiotic 3-chloroaniline disturbance favors deterministic assembly with a shift in function and structure of bacterial communities in sludge bioreactors

Authors: Ezequiel Santillan[1,2§], Hari Seshan[1,2,†§], and Stefan Wuertz[1,2,4*]

Affiliations:

[1]Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, 637551, Singapore.

[2]Department of Civil and Environmental Engineering, University of California, Davis, CA 95616, U.S.A.

[3]School of Civil and Environmental Engineering, Nanyang Technological University, Singapore, 639798, Singapore.

*Correspondence to: Stefan Wuertz, swuertz@ntu.edu.sg

[§]E.S. and H.S. contributed equally.

[†]Current affiliation: Stantec Australia Pty Ltd, 52 Merivale St, South Brisbane QLD 4101, Australia.

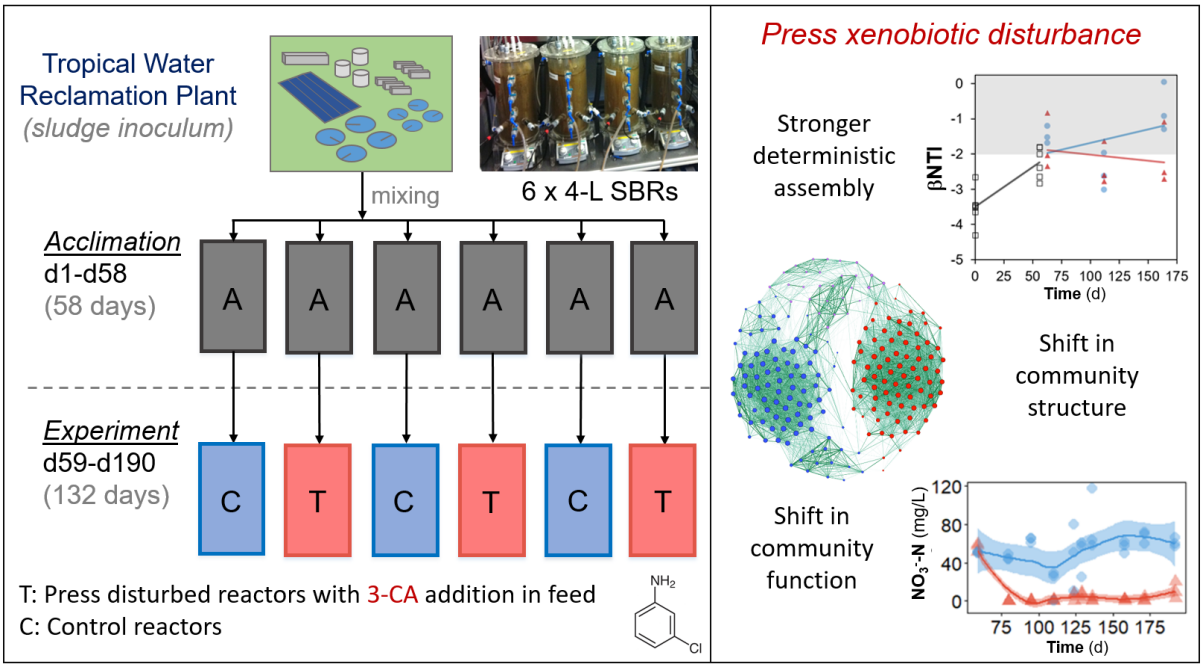
Abstract

Disturbance is thought to affect community assembly mechanisms, which in turn shape community structure and the overall function of the ecosystem. Continuous (press) disturbances can drive ecosystems to alternate stable states of community function and structure, but their effects on assembly mechanisms are still largely unknown. Here, we tested the effect of a press xenobiotic disturbance on the function, structure, and assembly of bacterial communities within a wastewater treatment system. Two sets of four-liter sequencing batch reactors were operated in triplicate with and without the addition of 3-chloroaniline for a period of 132 days, following 58 days of acclimation after inoculation with sludge from a full-scale treatment plant. Temporal dynamics of bacterial community structure were derived from 16S rRNA gene amplicon sequencing. Community function, structure and assembly differed between press disturbed and undisturbed reactors. Temporal partitioning of assembly mechanisms via phylogenetic and taxonomic null modelling analyses revealed that deterministic assembly prevailed for disturbed bioreactors, while the role of stochastic assembly was stronger for undisturbed reactors. Our findings are relevant because research spanning various disturbance types, environments and spatiotemporal scales is needed for a comprehensive understanding of the effects of press disturbances on assembly mechanisms, structure, and function of microbial communities.

Keywords: diversity, disturbance, community structure, stochastic assembly, deterministic assembly.

Synopsis: Continuous 3-chloroaniline addition promotes selective assembly mechanisms that alter the performance and composition of bacterial communities in wastewater treatment bioreactors.

Graphical abstract



Introduction

Sludge bioreactors for wastewater treatment are good model systems for microbial ecology studies[1]. These engineered systems involve measurable functions such as the removal of carbon and ammonia that are relevant in practice[2] and entail complex microbial communities in a controlled environment[3]. In nature, microbial communities drive all biogeochemical cycles while providing ecosystem functions necessary for all other life forms to exist[4]. The structure of these communities, which can be described in terms of α- and β-diversity metrics, is believed to shape the ecosystem function they provide[5]. Predicting and managing the functions of microbial communities based on a fundamental understanding of their relationship with community structure remains a key challenge for complex microbial systems.

Disturbance is thought to have direct impacts on ecosystems by shifting community structure and function[6]. Chloroanilines are intermediate breakdown products from the use of herbicides and pesticides in agriculture[7], which can be found in rubber, dye, polymer and pharmaceutical industrial wastewater treatment plants[8,9], but also in soil and natural aqueous environments[10]. These xenobiotic compounds are known to hamper both carbon and nitrogen removal in bioreactors for wastewater treatment[11], and can represent a disturbance in sludge bioreactor systems. Press disturbances, which inflict a long-term continuous alteration of taxa abundances by changing their environment[12], are relevant for both microbial ecology and environmental biotechnology as they can lead systems to alternative stable states with different community function and structure[13]. A rigorous evaluation of such variability requires designs that include replication at both press disturbed and undisturbed levels[14], yet this is not usually the case in studies employing sludge bioreactors.

Community assembly processes are fundamentally related to ecosystem function, as these are thought to shape community structure[15]. These processes can be either deterministic or stochastic, often acting in combination to shape patterns of community assembly[16-18]. In ecology, the contribution of assembly mechanisms is usually quantified via null model analyses[19]. Although disturbance is thought to be a main driver of these underlying community assembly mechanisms[20], a predictive understanding of its effects remains elusive[21]. Disturbance can favour stochastic assembly mechanisms that could lead communities to deviating states of function and structure[22,23]; hence, assessment of its effects also

demands replicated study designs[14,24]. Further, while several studies have reported patterns of community assembly in engineered microbial systems[25-28], relatively few have addressed the effects of disturbance on community assembly, structure and function of such systems, particularly for bioreactors treating wastewater.

The objective of this work was to test the effect of a press disturbance on community assembly mechanisms by introducing a xenobiotic compound, 3-chloroaniline (3-CA), in a replicated set of activated sludge bioreactors with a working volume of four liters, representing a mesocosm scale. Based on our findings in a prior reactor study at a microcosm scale[22], we expected to see a stronger deterministic effect at the disturbed level, together with an alteration in structure and function of bacterial communities in sludge bioreactors. Samples were analysed using 16S rRNA gene amplicon sequencing and effluent chemical characterization. Patterns of α- and β-diversity were employed to assess temporal dynamics of community structure. Assembly mechanisms were quantified via two different mathematical null models, one that assessed the phylogenetic turnover for each bioreactor, and another that evaluated the effective bacterial turnover expressed as a proportion of total bacterial diversity[29] across all replicate reactors.

Materials and Methods

Experimental design

Six sequencing batch bioreactors (SBRs) were operated in parallel and fed synthetic wastewater for an acclimation period of 58 d. The inoculum was taken from the aeration tanks of a full-scale wastewater treatment plant (WWTP) in Singapore. On day 0 of acclimation, the freshly sourced inoculum sludge was mixed well in the laboratory and distributed to all six reactors (2 L each). Each was topped up with 2 L of synthetic wastewater, to a working volume of 4 L per reactor. The reactors were run under identical conditions with 12-h cycles as follows: 20 min of feeding, 180 min of anoxic mixing, 440 min of aeration and mixing (dissolved oxygen, DO, maintained at 1-2 mg/L using a feedback loop where aeration would commence at 1 L air/min when probes measured DO as below 1 mg/L and aeration would stop when the DO reading was above 2 mg/L), 50 min of settling and 30 min of effluent (supernatant) discharge. Two liters of effluent were discharged at the end of every cycle and replaced

with 2 L of synthetic wastewater at the beginning of the next 12-h cycle, resulting in a hydraulic retention time (HRT) of 24 h. The mixed liquor temperature was maintained at 30°C using water jackets around the reactors and a re-circulating water heater. Solids were removed regularly from the mixed liquor to maintain a solids retention time (SRT) of about 30 d in each reactor.

The synthetic wastewater fed to all six reactors during the acclimation period was adapted from Hesselmann *et al.*[30] and contained the following (mg/L in each reactor after feeding and being diluted in mixed liquor): sodium acetate (112.5), dextrose (45), yeast extract (67.5), soy peptone (60), meat peptone (60), casein peptone (90), urea (15), ammonium bicarbonate (90), ammonium chloride (169), disodium hydrogen phosphate (720), potassium dihydrogen phosphate (130), calcium chloride dihydrate (10.5) and magnesium sulphate heptahydrate (112.5). The medium also contained 2 mL of the unaltered trace element stock[30] per liter of medium. The first six components contributed to about 500 mg COD/L of organic carbon concentration, while the next three (ammonium-based) components contributed to a nitrogen concentration of about 70 mg NH_4^+-N/L, both at the beginning of a cycle (*i.e.*, in the mixed liquor after feeding). Phosphates were used to buffer the medium and maintain a pH of around 7.5 to facilitate the nitrification process. Since the influent was diluted by a factor of 2 when fed to the mixed liquor, it should be noted that the concentration of each of the above constituents in undiluted influent (before feeding) was twice the concentration shown above. After the 58-d acclimation period, the sustained 3-CA input experiment was started and continued for 132 d. At the start of this experiment, three of the acclimated reactors were randomly assigned to the treatment group (press disturbed with 3-CA) and the other three were assigned to the control group (no 3-CA addition, or undisturbed). The cycle conditions and other parameters were kept the same among all six reactors and were identical to the conditions in the acclimation phase. The 3-CA inputs to the treatment reactors started on day 59. The medium used to feed the treatment reactors was slightly altered: the organic constituents (sodium acetate, dextrose, yeast extract, soy peptone, meat peptone and casein peptone) were scaled down by 20%, resulting in a total COD of 400 mg/L from these constituents. The remaining 20% of the COD was fed in the form of 3-CA, resulting in a mixed liquor 3-CA concentration of about 70 mg/L, a relevant level for WWTPs treating industrial or a mix of urban-industrial wastewaters[11]. The

three control reactors continued to receive the same 3-CA-free medium that was used during acclimation.

To assess the changes in community composition, samples of sludge were taken from all six reactors at five distinct time points, representing five different periods in the progression of the experiment. These five time points were day 0 (the day the reactors were inoculated for acclimation), day 56 (shortly before 3-CA addition to treatment reactors was commenced), day 63 (shortly after 3-CA addition to treatment reactors was commenced), day 112 (after process performance had stabilised) and day 164 (toward the end of the experiment). On each of these sampling days, sludge was collected directly from the mixed liquor at the beginning of a given anoxic phase, aliquoted into 1 mL in cryogenic tubes, flash-frozen in liquid nitrogen and stored at -80°C for subsequent molecular analysis via 16S rRNA gene amplicon sequencing and terminal restriction fragment length polymorphism (T-RFLP) analysis. The T-RFLP results are discussed elsewhere[31].

Water chemical analysis

Water quality parameters were measured using standard nutrient analysis techniques in accordance with standard methods[32] and targeted COD (standard methods 5220 D), nitrogen species (ammonium, nitrite and nitrate ions) using ion chromatography (standard methods 4500-NH3 for ammonium; 4110 B for nitrate and nitrite). Nitrogen species were also measured using spectrophotometric tests (Hach) to complement ion chromatography results. 3-CA was measured on a Shimadzu Prominence high pressure liquid chromatography system (Shimadzu) equipped with a UV-VIS PDA detector using an Ascentis C18 5-µm column (Sigma-Aldrich). An isocratic 50:50 Water:Methanol solvent was used at a flow of 0.3 mL/min, and 3-CA peaks were identified and measured at 199, 237 and 286 nm.

16S rRNA gene amplicon sequencing and reads processing

Genomic DNA was extracted from samples collected from bioreactors on the five days identified above (500 µL of sludge per sample) using the FastDNA Spin Kit for Soil (MP Biomedicals) with alterations to the manufacturer's protocol to increase DNA yield as detailed elsewhere[22]. Extracted DNA was amplified in triplicate 50-µL PCR reactions using primer set 530f/U1053r, which targets the V3~V4-V5 variable regions of the bacterial 16S rRNA gene[33]. Each 50-µL PCR reaction contained 25 µL of

ImmoMix reagent (Meridian Bioscience) and about 200 ng of DNA extract and nuclease-free water. The PCR program included an initial denaturation step at 95° C for 10 min, followed by 30 cycles of denaturation (95°C, 1 min), annealing (58°C, 30 s) and extension (72°C, 1 min). A final extension was carried out at 72° C for 7 min. The triplicate PCR amplicons from each sample-primer set combination were then pooled and purified using the QIAquick PCR purification kit (Qiagen), with a slightly altered protocol in which the final elution was performed on nuclease-free water pre-heated to 55°C and incubated for 5 min at 55° C before the final centrifugation. This modification was found to increase the final DNA yield. The purified amplicons were then inspected for quality using agarose gels (2% in TAE buffer) and quantified using a Qubit 3.0 fluorometer (ThermoFisher Scientific).

The libraries were sequenced in-house at SCELSE on an Illumina MiSeq (v.3) with 20% PhiX spike-in, at 300 bp paired-end read-length. Sequenced sample libraries were processed with the *dada2* (v.1.3.3) R-package[34], allowing inference of amplicon sequence variants (ASVs)[35], using R software[36] (v.3.6.3). Illumina adaptors and PCR primers were trimmed prior to quality filtering. Sequences were truncated after 280 and 255 nucleotides for forward and reverse reads, respectively. After truncation, reads with expected error rates higher than 3 and 5 for forward and reverse reads, respectively, were removed. Since paired-end reads could not be merged with a minimum overlap of 20 bp, only forward reads were kept for further analysis. Chimeric sequences (0.3% on average) were identified and removed. For a total of 30 samples, an average of 48,659 reads were kept per sample after processing, representing 61% of the average forward input reads. Taxonomy was assigned using the SILVA database (v.132)[37]. Samples were rarefied to the lowest number of reads (18,353) in a sample after processing (Fig. S1).

Bacterial community analysis and statistics

All reported p-values for statistical tests in this study were corrected for multiple comparisons using a False Discovery Rate (FDR) of 5%[38]. Community structure was assessed by a combination of principal coordinate analysis (PCoA) ordination of weighed Unifrac dissimilarity matrixes, constructed from Hellinger transformed abundance data using the *phyloseq*[39] R-package (v.1.30.0) in R. Differences in community structure across control and treatment reactors were assessed using multivariate tests of

permutational analysis of variance (PERMANOVA) and permutational analysis of dispersion (PERMDISP) on Bray-Curtis dissimilarity matrixes, constructed from square-root transformed abundance data using PRIMER (v.7)[40]. Hill diversity indices[41] were used to quantify α-diversity as described elsewhere[22]. Local polynomial regression fitting was applied using the *loess* function from the *ggplot2* package (v.3.3.2) in R[42], including 95% confidence intervals. Welch's t-test (two-tailed) was used for univariate testing. Heat maps for bacterial genera relative abundances were constructed using the *ampvis2* package (v.2.6.2) in R[43]. Composition-aware network analysis was performed on data from day 164 with Gephi (v.0.9.2) using SparCC correlation matrixes for the top 200 ASVs. Input correlation matrixes were generated using the *sparcc* function of the *SpiecEasi* (v.1.1.0) R package[44]. Non-strong correlations ($r_{adj} < 0.5$) were filtered out. Node clusters were defined by modularity class calculated using the Louvain method, and were coloured through identification of representative nodes at undisturbed and disturbed levels. Layout was adjusted using the Fruchterman Reingold method with default parameters. Node size was adjusted by degree, while edge thickness was adjusted by correlation strength.

Null model analyses on diversity

Further, community assembly processes were assessed from ASV data, thus reducing diversity estimation bias by obtaining 10-100 times fewer spurious units than with traditional OTU clustering[34,35]. The effect of underlying assembly mechanisms was first assessed using a phylogenetic-based null modelling approach. The model uses the β-mean nearest taxon distance (βMNTD)[45] which quantifies the phylogenetic distance between each ASV in one community, as a measure of the clustering of closely related ASVs. Phylogenetic relatedness of ASVs was characterized by multiple-alignment of ASV sequences using *decipher* (v.2.14.0) R-package[46]. The phylogenetic tree was then constructed and a GTR+G+I maximum likelihood tree was fitted using the *phangorn* (v.2.5.5) R-package[47]. To quantify the degree to which βMNTD deviates from a null model expectation, ASVs and abundances were shuffled across the tips of the phylogenetic tree. After shuffling, βMNTD was recalculated to obtain a null value, and repeating the shuffling 1,000 times provided a null distribution. The difference between observed βMNTD and the mean of the null distribution was measured in units of standard deviation,

which is referred to as the β-nearest taxon index (βNTI)[48]. A value of |βNTI| > 2 indicates that the observed turnover between a pair of communities is significantly deterministic, while |βNTI| < 2 suggests stochastic assembly[17]. This analysis was done using the *phylocomr* (v.0.3.2) R-package[49]. To test for a phylogenetic signal across phylogenetic distances, Mantel correlograms were constructed relating between-ASV niche differences to between-ASV phylogenetic distances across a given phylogenetic distance, following the previously described methodology[17,48]. Environmental niches were constructed from bioreactor effluent data (soluble chemical oxygen demand, 3-chloroaniline, ammonium, nitrite and nitrate as nitrogen). Phylogenetic distances were quantified for 50 phylogenetic distance bins, while the significance of Pearson correlations was assessed using 1,000 permutations and FDR (5%) correction.

We further assessed assembly mechanisms by using an alternative null modelling methodology, which assumes that species interactions are not important for community assembly[19] and quantifies the effective bacterial turnover expressed as a proportion of total bacterial diversity. It was developed for woody plants[50] and recently applied to sludge[18,22] and groundwater[28] microbial communities. The model defines β-diversity as the β-partition ($\beta = 1 - \bar{\alpha}/\gamma$). To adapt it to handle microbial community data, we considered 'species' in the model as ASVs, while each individual count was one read within the corresponding dataset. The model randomizes the location of each individual within the independent replicate reactors for each of the control and treatment levels, while maintaining the total quantity of individuals per reactor, the relative abundance of each 'species' (*i.e.,* ASVs), and the γ-diversity. This way it takes into account both composition and relative abundances. We applied the model across different time points of the experiment. Control and treatment levels had three replicate reactors each, thus acclimation phase data were assigned to the same two groups of three reactors each for model estimations. Each step of the null model calculated expected mean α-diversities per treatment level and then estimated an expected β-partition. After 10,000 repetitions, the means of the distribution of random β-partitions ($\overline{\beta_{\mathrm{exp}}}$) for each treatment level were calculated. The relative contribution of deterministic assembly mechanisms was then quantified using the stochastic intensity (SI) metric, which measured the deviation of the observed β-diversity compared to that expected by chance as follows: $SI = 1 -$

$\left|\beta_{obs} - \overline{\beta_{exp}}\right| / \beta_{obs}$. SI is the complement of the deterministic strength (DS) metric ($DS = \left|\beta_{obs} - \overline{\beta_{exp}}\right| / \beta_{obs}$), defined previously[18,22]. Higher values of SI indicated a lower deviation of the observed β-diversity from the null β-diversity expectation, thus suggesting a stronger effect of stochastic-based mechanisms. Contrarily, lower SI values indicated a bigger difference between observed and null β-diversities, suggesting a more important role of deterministic mechanisms of assembly.

Results

Dynamics of process performance

There was a clear distinction in process performance (*i.e.*, ecosystem function) between control and treatment levels. Nitrate was found in the effluent of control reactors (Fig. 1C), which together with low levels of residual ammonia and nitrite (Fig. 1A-B) indicated complete nitrification. Conversely, treatment reactors initially displayed higher COD, $NH_4^+ - N$ and $NO_2^- - N$ concentrations in the effluent due to inhibition of nitrification (Fig. 1, Fig. S2). However, during the course of the experiment the microbial communities in the three treatment reactors acquired the ability to metabolize 3-CA (Fig. 1E) and perform partial nitrification (ammonia oxidation) (Fig. 1B). The metabolism of 3-CA in treatment reactors occurred within the first three weeks of 3-CA input and was followed by a gradual recovery of ammonia oxidation several weeks later. The lack of nitrification in treatment reactors is evidenced by the lower levels of ammonia removal (Fig. S2A). The percent removal values in Fig. S2A are based on an ammonia concentration of 70 mg/L at the beginning of a cycle (*i.e.*, after dilution in mixed liquor). Subsequent accumulation of ammonia due to the lack of nitrification in treatment reactors led to negative percent removal values (Fig. S2A), since the influent concentration of ammonia-nitrogen was 140 mg/L. The partial recovery of ammonia oxidation in treatment reactors was seen toward the end of the experiment, and is evidenced by the slight fall in effluent ammonia concentrations in treatment reactors (Fig. 1A), accompanied by the slight rise in nitrite over the course of the experiment (Fig. 1B). The percent removal plots did not show this recovery in ammonia oxidation as clearly because of the accumulation in reactors over most of the experimental period. However, the general rise in percent removal toward the end of the experiment, which coincided with the rise in

effluent nitrite concentration, indicates the recovery of ammonia oxidation. A more detailed analysis of changes in process performance under 3-CA disturbance in this experiment was the focus of an additional study[31]. In that work, analysis of process variables are explored in full SBR cycles during the course of the experiment, together with an assessment of the metabolic mechanisms behind 3-CA removal, for a deeper understanding of the effect of 3-CA disturbance on activated sludge performance.

Dynamics of bacterial community structure

Throughout the acclimation phase there were significant changes in bacterial community structure, both in terms of α-diversity (Fig. 2) ($P\text{-}_{\text{Welch's t-test}} < 0.009$) and β-diversity (Fig. 3) ($P\text{-}_{\text{PERMANOVA}} < 0.0001$), from the initial wastewater treatment plant inoculum (day 0) all the way to the acclimated sludge on day 58. α-diversity after the acclimation period varied over time, but did not differ significantly between control and treatment reactors ($P\text{-}_{\text{Welch's t-test}} > 0.10$), although control reactors displayed higher α-diversity only towards the end of the study (Fig. 2A-C). However, temporal patterns of β-diversity showed bacterial communities clustering separately in control and treatment reactors (Fig. 3). Such differentiation was statistically significant from day 112 onwards ($P\text{-}_{\text{PERMANOVA}} < 0.016$), with no effects of heteroscedasticity ($P\text{-}_{\text{PERMDISP}} > 0.41$, Table S1). Bacterial succession was also evident from the analysis of relative abundances of specific taxa, which were different at each stage of the study (Fig. 4A). Genera like *Propioniciclava*, *OLB8* and *Terrimonas* prevailed initially on day 0. After acclimation *Paracoccus*, *Kouleothrix* and *Alicycliphilus* dominated across control reactors, while *Ca.* Competibacter, *Dokdonella* and *Gemmatimonas* prevailed in treatment reactors. Some taxa like *Tetrasphaera* and *Microlunatus* increased in relative abundance for all reactors throughout the study. Composition-aware correlation network analysis on day 164 via SparCC, also displayed separate clusters of bacterial ASVs based on node modularity (Fig. 4B). See Seshan *et al.*[31] for a detailed analysis of the effect of a 3-CA press disturbance on nitrifying genera and the putative microorganisms involved in 3-CA degradation.

Dynamics of bacterial community assembly mechanisms

Two different approaches were applied to quantify assembly mechanisms, based on the deviation of observed β-diversity from the expected β-diversity from null modelling. Phylogenetic turnover was first assessed using the β-nearest taxon index (βNTI). Treatment reactors had mostly mean values of βNTI < -2, indicating deterministic assembly (Fig. 5A). For control reactors, values of |βNTI| < 2 indicated a significant effect of stochastic assembly (Fig. 5A). The observed βNTI patterns were not driven by a particular reactor (Fig. S3A). A significant phylogenetic signal was found across relatively short phylogenetic distances via mantel correlogram analysis (Fig. S3B), which supported the use of βNTI. Concurrently, the effective bacterial turnover expressed as a proportion of total bacterial diversity was quantified via the stochastic intensity (SI) metric. Similarly, SI values trended higher in control reactors, suggesting a greater relative role of stochastic assembly mechanisms compared to treatment reactors (Fig. 5B).

Discussion

For this study, we hypothesized that press disturbance by a xenobiotic would alter function, structure and assembly of sludge bioreactor bacterial communities. Chloroanilines are toxic substances that can easily diffuse within the natural environment and are difficult to remediate, posing an environmental problem for natural waters and soils where their concentrations are increasing[10]. This work used 3-CA as disturbance, a compound known to have a detrimental effect on relevant functions within wastewater bacterial communities[11,22]. Indeed, function (in terms of process performance) and structure (in terms of β-diversity) clearly differed between control and treatment reactors. Disturbance impaired nitrification in treatment reactors (Fig. 1B-C), coinciding with a marked community differentiation from the undisturbed control reactors (Figs. 3, 4), although only a few taxa are known to participate in autotrophic nitrification. Community evenness, a metric of α-diversity, was suggested to be a key factor in preserving the functional stability of an ecosystem[51]. Therefore, control reactors that displayed better COD removal and complete nitrification with almost no residual $NH_4^+ - N$ or $NO_2^- - N$, were expected to harbor more diverse communities than treatment reactors. However, we found no significant differences in α-diversity (Fig. 2) in terms of number of ASVs (^{0}D) and in how evenly the

relative abundances were distributed among taxa in control and treatment reactors (^{1}D, ^{2}D). The two main genera in both control and treatment reactors (Fig. 3) were *Tetrasphaera* and *Microlunatus*, which include polyphosphate accumulating organisms (PAO) in activated sludge systems. These taxa likely benefited from the high amount of PO_4^{3-}-P in the synthetic wastewater and the alternating anoxic/anaerobic-aerobic cycle conditions in bioreactors. In control reactors, complete nitrification was achieved by the end of an SBR cycle, meaning that after discharging half the reactor volume there was enough nitrate present to support denitrifying organisms like *Paracoccus* and *Alicycliphilus*, which could use it as terminal electron acceptor during the anoxic phase of the cycle. The prevailing nitrifying genera, *Nitrosomonas* and *Nitrospira*, were relatively less abundant in treatment reactors than in control reactors, coincident with a reduction in ammonia and nitrite oxidation activity (details in Seshan *et al.*[31]). On the other hand, organisms like *Candidatus* Competibacter (a glycogen accumulating organism) and *Gemmatimonas* (a PAO) prevailed in treatment reactors, as was also reported for microcosm sludge bioreactors under a similar 3-CA pollutant disturbance[52]. The observed differences in function and dominance of taxa under press disturbed and undisturbed conditions are reasonable, given that organisms face tradeoffs when utilizing resources to maximize their fitness depending on the interactions within their habitat[52,53].

Community assembly mechanisms were also different between control and treatment reactors, coherent with the idea that underlying assembly mechanisms shape community function and structure[15]. In our study, disturbance favoured deterministic assembly over time, regardless of the type of null model approach employed for quantification. Phylogenetic-based assembly analysis showed a significant effect of stochastic assembly for control reactors, while treatment reactors displayed dominance of deterministic processes categorized as homogeneous selection[48] (Fig. 5A). In terms of bacterial turnover assessed via the relative stochastic intensity metric, disturbed reactors showed a stronger role of deterministic mechanisms (Fig. 5B), likely due to the selective pressure via environmental filtering[54] that occurs under disturbance. Further, β-diversity was always greater than expected for bacterial taxa under the taxonomic null model (Table S2), implying that organisms tended to be more aggregated within replicate bioreactors than expected by chance. Aggregation can be explained by processes of dispersal limitation[55], due to operating reactors as closed systems without immigration as part of the

experimental design of this study, and by habitat filtering[56], owing to the regular conditions at undisturbed and disturbed levels. Observing similar trends for the effects of disturbance on community assembly via both methods employed is not trivial, as the results from null model assessments are dependent on the algorithms, models and diversity metrics used[57]. The βNTI approach has the advantage of including phylogeny in the analysis[17,48], but it does not take into account all independent replicates for a given time point to build the null model distribution like the effective bacterial turnover null modelling approach[22,50]. Therefore, the combination of both methodologies allowed us to better exploit the replicated design of this study. Further, differences in community assembly mechanisms operating at phylogenetic and taxonomic levels could be the reason why the acclimation phase showed different trends in assembly mechanisms for each of these methodologies employed.

Our findings concur with those of previous studies of press disturbance in sludge bioreactors using the same micro pollutant (3-CA) at different spatial and temporal scales, as well as different methodologies to assess community structure and assembly. A mesocosm study using 2-L membrane bioreactors, operated during 70 days following a 6-month acclimation period, reported higher similarity among disturbed reactors than control reactors as measured by T-RFLP analysis of a 16S rRNA gene amplicon[11], although without quantifying assembly mechanisms. Observations derived from community dissimilarities (*i.e.*, patterns of β-diversity) can highlight deterministic effects but have low power to infer stochasticity[24], the reason why it is important to partition community assembly mechanisms via null model analyses. A recent study quantified assembly mechanisms via null model analysis on metagenomics genus-level data[22], finding higher stochastic intensity for sludge microcosm reactors that were undisturbed compared to those that were press disturbed with 3-CA. The authors employed a bioreactor scale (20 mL) two orders of magnitude smaller than this work (4 L) and employed a shorter experimental period (35 d) without acclimation. Diversity is multidimensional and scale-dependent[58], thus assembly mechanisms and the results of null model analyses can differ at different spatial and temporal scales[21,57]. Hence, the observed similar effects of press disturbance in community assembly found at different temporal and volume (*i.e.*, microcosm and mesocosm) scales are relevant. Disturbance was also found to promote deterministic mechanisms in a mesocosm study

that used a similar replicated bioreactor system of 5-L SBRs operated for 127 d (including 53 d of acclimation) and the same combination of null-model approaches to quantify assembly mechanisms, but employed a change in organic loading as press disturbance[18], instead of a xenobiotic. These are relevant comparisons given the multidimensional nature of disturbances, which can be of different types and occur through space and time at different frequencies and intensities and to varying extents[59,60]. Hence, our findings in this study strengthen the utility of exploring the role of disturbance in stochastic and deterministic assembly mechanisms for microbial communities.

Conclusions

Overall, this study advances our understanding of the response of complex microbial systems to a xenobiotic compound, by using a joint evaluation of assembly mechanisms, community structure and function of bacterial taxa. In our replicated mesocosm bioreactor system, disturbance altered community function and structure and elicited stronger deterministic mechanisms of community assembly as assessed by two different null model approaches. While this study employed the addition of an aromatic pollutant in the bioreactor feed as a type of press disturbance, more research is needed on different types of disturbances (*e.g.,* pH changes, invading taxa and temperature variations) within different complex microbial systems to further test the general validity of our observations. In this manner, studies covering different temporal and spatial scales, environments and types of disturbance could lead to a general understanding of how press disturbances shift the function, structure, and assembly mechanisms of microbial communities.

Acknowledgements

This research was supported by the Singapore National Research Foundation and Ministry of Education under the Research Centre of Excellence Program. We thank DI Drautz-Moses for her support with the 16S rRNA gene amplicon library preparation and sequencing pipelines employed.

Author Contributions

HS designed the experiment and SW obtained the funding for the study. HS performed the experiments. HS and ES performed the molecular work in preparation for amplicon sequencing. ES did the

bioinformatics and null model analyses. ES interpreted the data and elaborated the main arguments in the manuscript. All authors contributed to manuscript writing and editing.

Data availability

DNA sequencing data are available at NCBI BioProject under accession no. PRJNA720804. See supporting information for details on temporal influent removal percentages at the end of a reactor cycle, null model and multivariate analyses, and 16S rRNA gene data rarefaction. R-script and source data to run the taxonomic null model analysis are provided as supplementary files. This material is available free of charge via the internet at http://pubs.acs.org.

Competing interests

The authors declare no competing interests.

References

1 Daims, H., Taylor, M. W. & Wagner, M. Wastewater treatment: a model system for microbial ecology. *Trends Biotechnol.* **2006**, 24 (11), 483-489.

2 Timmis, K. *et al.* The contribution of microbial biotechnology to sustainable development goals. *Microb. Biotechnol.* **2017**, 10 (5), 984-987.

3 Seviour, R. & Nielsen, P. *Microbial Ecology of Activated Sludge*. (IWA Publishing, London, UK, **2010**).

4 Flemming, H.-C. & Wuertz, S. Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.* **2019**, 17, 247–260.

5 Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **2012**, 486 (7402), 207-214.

6 Cain, M., Bowman, W. & Hacker, S. *Ecology*. 3 ed. (Sinauer Associates Inc., Sunderland, Massachusetts, **2014**).

7 Verhagen, P., Destino, C., Boon, N. & De Gelder, L. Spatial heterogeneity in degradation characteristics and microbial community composition of pesticide biopurification systems. *J. Appl. Microbiol.* **2015**, 118 (2), 368-378.

8 Boon, N., Top, E. M., Verstraete, W. & Siciliano, S. D. Bioaugmentation as a tool to protect the structure and function of an activated-sludge microbial community against a 3-chloroaniline shock load. *Appl. Environ. Microbiol.* **2003**, 69 (3), 1511-1520.

9 Jen, J. F., Chang, C. T. & Yang, T. C. On-line microdialysis-high-performance liquid chromatographic determination of aniline and 2-chloroaniline in polymer industrial wastewater. *J. Chromatogr. A* **2001**, 930 (1-2), 119-125.

10 Tasca, A. L. & Fletcher, A. State of the art of the environmental behaviour and removal techniques of the endocrine disruptor 3,4-dichloroaniline. *J. Environ. Sci. Heal. A* **2017**, 0, 1-11.

11 Falk, M. W. & Wuertz, S. Effects of the toxin 3-chloroaniline at low concentrations on microbial community dynamics and membrane bioreactor performance. *Water Res.* **2010**, 44 (17), 5109-5115.

12 Shade, A. *et al.* Fundamentals of microbial community resistance and resilience. *Front. Microbiol.* **2012**, 3, 1-19.

13 Botton, S., van Heusden, M., Parsons, J. R., Smidt, H. & van Straalen, N. Resilience of microbial systems towards disturbances. *Crit. Rev. Microbiol.* **2006**, 32 (2), 101-112.

14 Prosser, J. I. Replicate or lie. *Environ. Microbiol.* **2010**, 12 (7), 1806-1810.

15 Leibold, M. A., Chase, J. M. & Ernest, S. K. M. Community assembly and the functioning of ecosystems: how metacommunity processes alter ecosystems attributes. *Ecology* **2017**, 98 (4), 909-919.

16 Stegen, J. C., Lin, X. J., Konopka, A. E. & Fredrickson, J. K. Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* **2012**, 6 (9), 1653-1664.

17 Dini-Andreote, F., Stegen, J. C., van Elsas, J. D. & Salles, J. F. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci. USA* **2015**, 112 (11), E1326-E1332.

18 Santillan, E., Constancias, F. & Wuertz, S. Press disturbance alters community structure and assembly mechanisms of bacterial taxa and functional genes in mesocosm-scale bioreactors. *mSystems* **2020**, 5 (4), e00471-00420.

19 Gotelli, N. J. & McGill, B. J. Null versus neutral models: what's the difference? *Ecography* **2006**, 29 (5), 793-800.

20 Ferrenberg, S. *et al.* Changes in assembly processes in soil bacterial communities following a wildfire disturbance. *ISME J.* **2013**, 7 (6), 1102-1111.

21 Nemergut, D. R. *et al.* Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* **2013**, 77 (3), 342-356.

22 Santillan, E., Seshan, H., Constancias, F., Drautz-Moses, D. I. & Wuertz, S. Frequency of disturbance alters diversity, function, and underlying assembly mechanisms of complex bacterial communities. *NPJ Biofilms Microbiomes* **2019**, 5 (1), 1-8.

23 Zhou, J. Z. *et al.* Stochastic assembly leads to alternative communities with distinct functions in a bioreactor microbial community. *mBio* **2013**, 4 (2), 1-8.

24 Zhou, J. & Ning, D. Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* **2017**, 81 (4), 1-32.

25 Liébana, R. *et al.* Combined deterministic and stochastic processes control microbial succession in replicate granular biofilm reactors. *Environ. Sci. Technol.* **2019**, 53 (9), 4912-4921.

26 Zhang, B. *et al.* The biogeography and assembly of microbial communities in wastewater treatment plants in China. *Environ. Sci. Technol.* **2020**, 54 (9), 5884-5892.

27 Wu, L. *et al.* Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nat. Microbiol.* **2019**, 1-13.

28 Zhou, J. Z. *et al.* Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proc. Natl. Acad. Sci. USA* **2014**, 111 (9), E836-E845.

29 Tuomisto, H. A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* **2010**, 33 (1), 2-22.

30 Hesselmann, R. P. X., Werlen, C., Hahn, D., van der Meer, J. R. & Zehnder, A. J. B. Enrichment, phylogenetic analysis and detection of a bacterium that performs enhanced biological phosphate removal in activated sludge. *Syst. Appl. Microbiol.* **1999**, 22 (3), 454-465.

31 Seshan, H. *et al.* Metagenomics and metatranscriptomics reveal pathway of 3-chloroaniline degradation in wastewater reactors. *bioRxiv* **2021**.

32 APHA-AWWA-WEF. *Standard methods for the examination of water and wastewater*. 22 ed. (AWWA, Washington D.C., USA, **2005**).

33 Thijs, S. *et al.* Comparative evaluation of four bacteria-specific primer pairs for 16S rRNA gene surveys. *Front. Microbiol.* **2017**, 8, 1-15.

34 Callahan, B. J. *et al.* DADA2: High resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, 13 (7), 581-583.

35 Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **2017**, 11, 2639–2643.

36 R: A language and environment for statistical computing v. 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria, 2020).

37 Glöckner, F. O. *et al.* 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J. Biotechnol.* **2017**, 261, 169-176.

38 Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. Ser. B. (Method.)* **1995**, 57 (1), 289-300.

39 McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **2013**, 8 (4), e61217.

40 Clarke, K. R. & Gorley, R. N. *PRIMER v7: User Manual/Tutorial*. (PRIMER-E, Plymouth, UK, **2015**).

41 Hill, M. O. Diversity and evenness: a unifiying notation and its consequences. *Ecology* **1973**, 54 (2), 427-432.

42 Wickham, H. *ggplot2: elegant graphics for data analysis*. 2nd ed. (Springer-Verlag, New York, **2016**).

43 Albertsen, M., Karst, S. M., Ziegler, A. S., Kirkegaard, R. H. & Nielsen, P. H. Back to basics - the influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. *PLoS One* **2015**, 10 (7), 15.

44 Kurtz, Z. D. *et al.* Sparse and Compositionally Robust Inference of Microbial Ecological Networks. *PLoS Comp. Biol.* **2015**, 11 (5), e1004226.

45 Webb, C. O., Ackerly, D. D., McPeek, M. A. & Donoghue, M. J. Phylogenies and Community Ecology. *Annu. Rev. Ecol. Syst.* **2002**, 33 (1), 475-505.

46 Wright, E. S. Using decipher v2.0 to analyze big biological sequence data in R. *R Journal* **2016**, 8 (1), 352-359.

47 Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **2010**, 27 (4), 592-593.

48 Stegen, J. C. *et al.* Quantifying community assembly processes and identifying features that impose them. *ISME J.* **2013**, 7 (11), 2069-2079.
49 Webb, C. O., Ackerly, D. D. & Kembel, S. W. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* **2008**, 24 (18), 2098-2100.
50 Kraft, N. J. B. *et al.* Disentangling the brivers of β diversity along latitudinal and elevational gradients. *Science* **2011**, 333 (6050), 1755-1758.
51 Wittebolle, L. *et al.* Initial community evenness favours functionality under selective stress. *Nature* **2009**, 458 (7238), 623-626.
52 Santillan, E., Seshan, H., Constancias, F. & Wuertz, S. Trait-based life-history strategies explain succession scenario for complex bacterial communities under varying disturbance. *Environ. Microbiol.* **2019**, 21 (10), 3751-3764.
53 Ho, A., Di Lonardo, D. P. & Bodelier, P. L. E. Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiol. Ecol.* **2017**, 93 (3), fix006-fix006.
54 Kraft, N. J. B. *et al.* Community assembly, coexistence and the environmental filtering metaphor. *Funct. Ecol.* **2015**, 29 (5), 592-599.
55 Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C. & Martiny, J. B. H. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* **2012**, 10 (7), 497-506.
56 Steinauer, K. *et al.* Plant diversity effects on soil microbial functions and enzymes are stronger than warming in a grassland experiment. *Ecology* **2015**, 96 (1), 99-112.
57 Gotelli, N. J. & Ulrich, W. Statistical challenges in null model analysis. *Oikos* **2012**, 121 (2), 171-180.
58 Chase, J. M. *et al.* Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. *Ecol. Lett.* **2018**, 21 (11), 1737-1751.
59 Miller, A. D., Roxburgh, S. H. & Shea, K. How frequency and intensity shape diversity-disturbance relationships. *Proc. Natl. Acad. Sci. USA* **2011**, 108 (14), 5643-5648.
60 Graham, E. B. *et al.* Toward a generalizable framework of disturbance ecology through crowdsourced science. *Frontiers in Ecology and Evolution* **2021**, 9 (76).

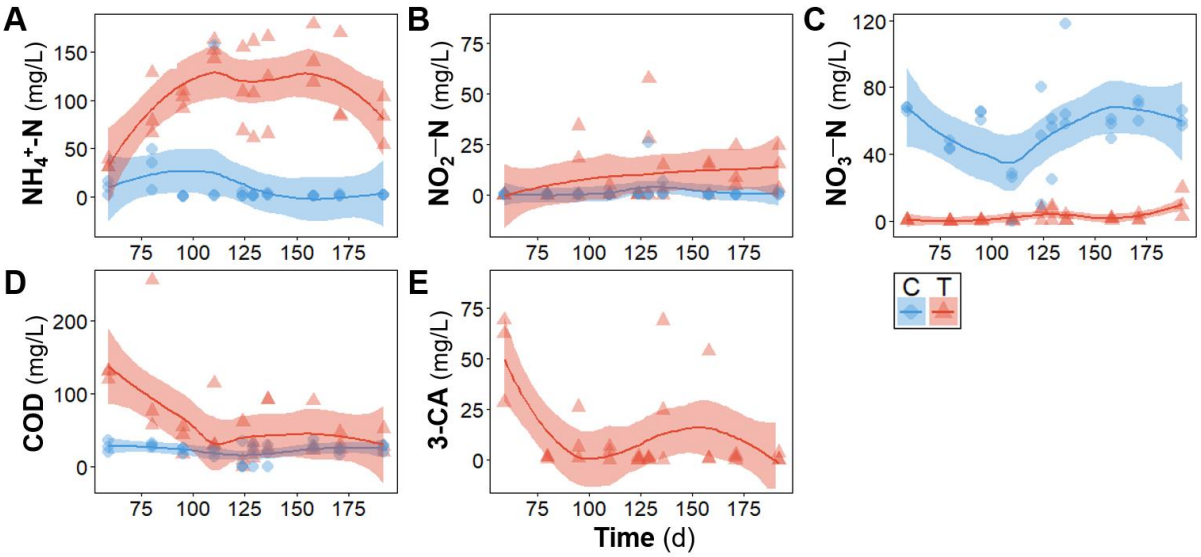
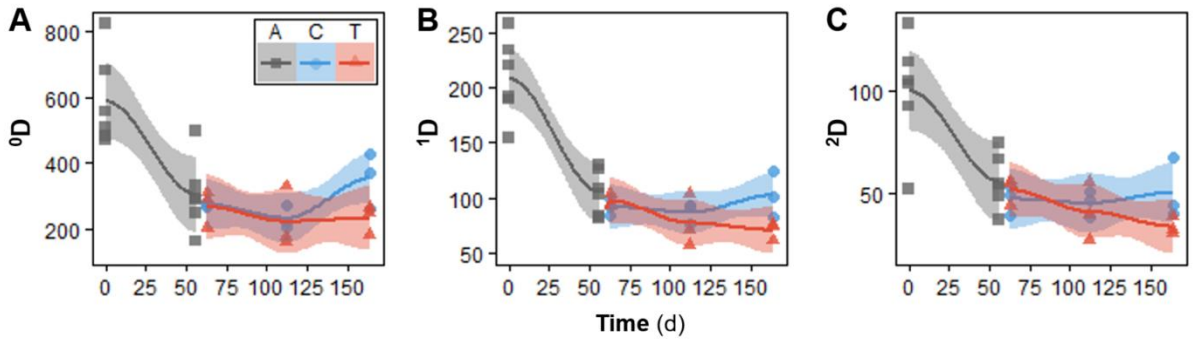
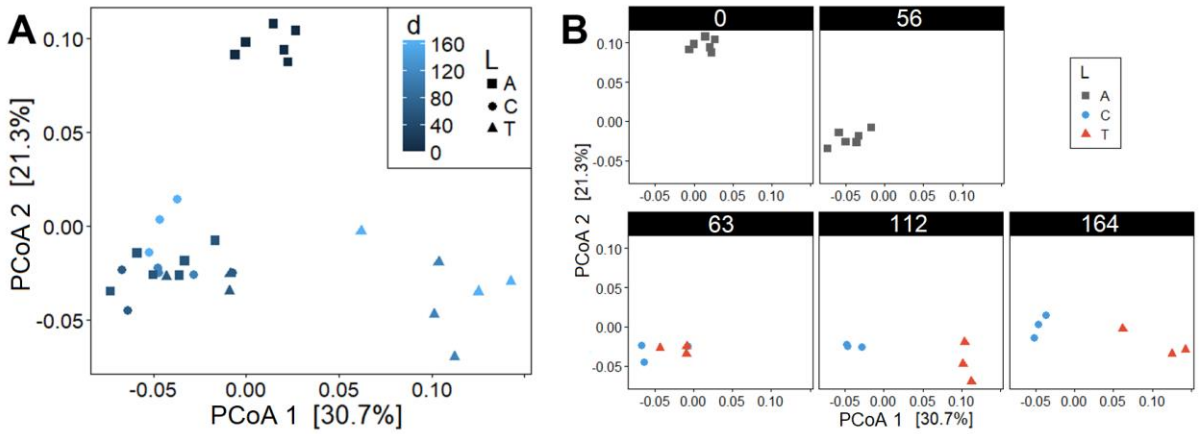


Figure 1 – Temporal effluent concentrations (mg/L) at the end of a reactor cycle, after the acclimation period (from d59 onwards). (**A**) Ammonium, (**B**) nitrite, and (**C**) nitrate as nitrogen, (**D**) organic carbon as soluble chemical oxygen demand (COD) and (**E**) 3-chloroaniline. Each point represents a different reactor for a given day. Reactor type: C, control (blue, n = 3); T, treatment (red, n = 3). Lines display polynomial regression fitting, while shaded areas represent 95% confidence intervals. No 3-CA was fed to control reactors, thus no blue points are displayed in panel **E**.



548

549 **Figure 2** - Temporal dynamics of true α-diversity Hill numbers for bacterial ASVs. Hill number orders:
550 (**A**) ^{0}D, (**B**) ^{1}D, and (**C**) ^{2}D. Each point represents a different reactor on a given day. Reactor type: A,
551 acclimation (grey, n = 6); C, control (blue, n = 3); T, treatment (red, n = 3). Lines refer to polynomial
552 regression fitting, while shaded areas represent 95% confidence intervals.



553

554 **Figure 3** - Temporal dynamics of community structure for bacterial ASVs evaluated through PCoA
555 ordination (weighed Unifrac β-diversity). Panels: (**A**) all time points (time increases from dark to light
556 blue) and (**B**) separate time points. Reactor type: A, acclimation (squares, n = 6); C, control (circles, n
557 = 3); T, treatment (triangles, n = 3). Each point represents a different reactor for a given day.

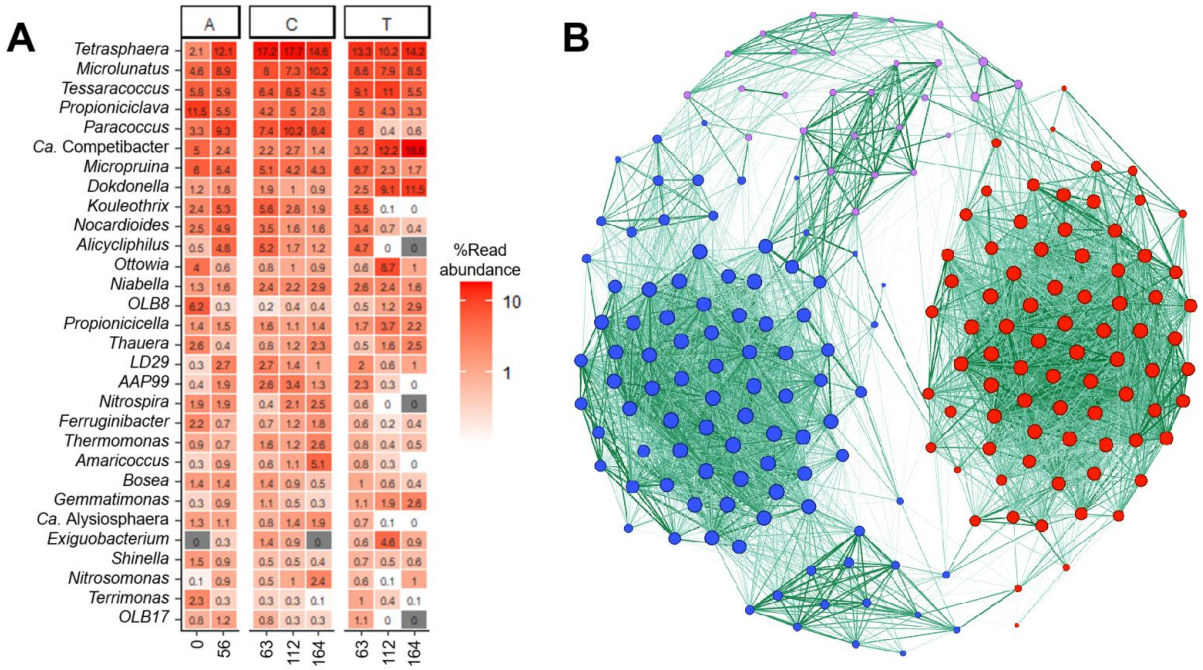


Figure 4 – (**A**) Community structure dynamics for bacterial genera, assessed through 16S rRNA gene amplicon sequencing. The 30 most abundant genera are shown. Columns represent the average percentage read abundance among reactors for a given reactor type and day. Reactor type: A, acclimation (left, n = 6); C, control (middle, n = 3); T, treatment (right, n = 3). (**B**) SparCC correlation network for the top 200 most abundant bacterial ASVs on d164. Clusters are coloured by modularity class, with blue nodes prevailing in control reactors (n = 3) and red nodes in treatment reactors (n = 3). For clarity, only strong SparCC correlations ($r_{adj} \geq 0.50$) are drawn in the plot. Edge thickness represents correlation strength and node size represents degree.

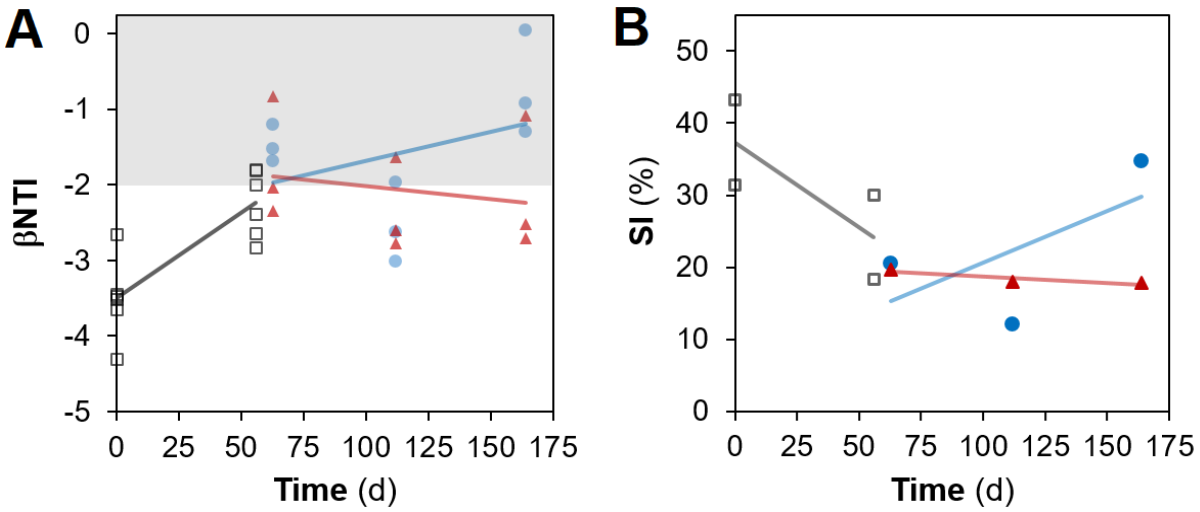


Figure 5 – Temporal dynamics of community assembly for bacterial ASVs evaluated via (**A**) nearest taxon index (βNTI) and (**B**) stochastic intensity (SI), from null model analyses. Reactor type: A, acclimation (grey, n = 6); C, control (blue, n = 3); T, treatment (red, n = 3). Lines represent linear regression fitting. Shaded in grey in (**A**) is the zone where stochastic processes significantly dominate |βNTI| < 2. Each point in (**B**) involves three replicates for SI calculation.