Title: Microbial residence time is a controlling parameter of the taxonomic composition and functional profile of microbial communities

1

2   **Running Title :** Microbial residence time effects on community

3

4   **Authors:** Cresten Mansfeldt[1*], Stefan Achermann[1,2*], Yujie Men[3], Jean-Claude Walser[4], Kris Villez[5],

5   Adriano Joss[5], David R. Johnson[6], Kathrin Fenner[1,2,7]

6

7   *These authors contributed equally to this work

8

9   1. Department of Environmental Chemistry, Swiss Federal Institute of Aquatic Science and Technology,

10   Eawag, Überlandstrasse 133, 8600 Dübendorf, Switzerland

11   2. Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, 8092 Zürich, Switzerland.

12   3. Department of Civil and Environmental Engineering, University of Illinois, 205 N. Mathews Ave.,

13   Urbana Illinois 61801, USA

14   4. Department of Environmental Systems Science, Genetic Diversity Centre, ETH Zürich,

15   Universitätstrasse 16, 8006 Zürich, Switzerland

16   5. Department of Process Engineering, Swiss Federal Institute of Aquatic Science and Technology,

17   Eawag, Überlandstrasse 133, 8600 Dübendorf, Switzerland

18   6. Department of Environmental Microbiology, Swiss Federal Institute of Aquatic Science and

19   Technology, Eawag, Überlandstrasse 133, 8600 Dübendorf, Switzerland

20   7. Department of Chemistry, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

21

24

Abstract:

A remaining challenge within microbial ecology is to understand the determinants of richness and diversity observed in environmental microbial communities. In a range of systems, including activated sludge bioreactors, the microbial residence time (MRT) has been previously shown to shape the microbial community composition. However, the physiological and ecological mechanisms driving this influence have remained unclear. Here, this relationship is explored by analyzing an activated sludge system fed with municipal wastewater. Using a model designed in this study based on Monod-growth kinetics, longer MRTs were shown to increase the range of growth parameters that enable persistence, resulting in increased richness and diversity in the modelled community. In laboratory experiments, six sequencing batch reactors treating domestic wastewater were operated in parallel at MRTs between 1-15 d. The communities were characterized using both 16S ribosomal RNA and non-target messenger RNA sequencing (metatranscriptomic analysis), and model-predicted monotonic increases in richness were confirmed in both profiles. Accordingly, taxonomic Shannon diversity also increased with MRT. In contrast, the diversity in enzyme class annotations resulting from the metatranscriptomic analysis displayed a non-monotonic trend over the MRT gradient. Disproportionately high abundances of transcripts encoding for rarer enzymes occur at longer MRTs and contribute to the disconnect between taxonomic and functional diversity profiles.

Introduction:

Environmental microbial communities often house a rich and diverse set of species and expressed enzymes[1,2]. A remaining challenge within microbial ecology is to understand the mechanisms driving the differences in metabolic and taxonomic diversity between communities[3]. Of the influencing mechanisms, the microbial residence time (MRT; the average amount of time a microorganism resides in a system) has been postulated to be a key parameter influencing microbial diversity[4,5]. Recent investigations in engineered systems showed that as MRT increases, the diversity and richness of the community increases as well. However, specific studies exploring the relationship between MRT and community composition

51 have shown opposing or more confounded trends[6,7] (albeit with differing experimental setups and analysis

52 methods), suggesting that the relationship between MRT of a system and community composition is

53 complex. The influence of MRT is also relevant in natural[8] and host-associated systems[9] suggesting that

54 more clearly identifying the influence of this parameter on community composition in engineered

55 environments may provide insights that are also relevant to other systems.

56

57   In addition to the influence on taxonomic diversity and composition, communities also express

58 more functions at longer MRT[4]. In wastewater treatment, functions related to substrate transformation

59 have been demonstrated to emerge at longer MRT, e.g., nitrification and the biotransformation of trace

60 organics[10,11]. In a survey of 10 wastewater treatment plants, functional richness was positively associated

61 with taxonomic richness, and both parameters were in turn positively associated with plant performance in

62 terms of trace organic contaminant removal[12]. By contrast, additional studies have noted that expressed

63 functional richness and diversity may not be directly related to taxonomic parameters[13-15]. In streams[16],

64 forests[17,18], and host associated communities[19], the monitored functional signals were independent of the

65 parameters controlling the taxonomic profiles. Both the taxonomic and functional profiles must be

66 monitored to understand further the linkage between community structure and function and to characterize

67 more accurately the influence of an external variable (such as MRT) on the community[20].

68

69   In this study, the influence of the MRT on the observed taxonomic composition and functional

70 profile of microbial communities cultivated in six parallel lab-scale sequencing batch reactors (SBR)

71 treating domestic wastewater was explored experimentally and described using a Monod-model.

72 Wastewater bioreactors provide a controllable experimental system[21] with available established

73 computational models[22-24] that have provided previous insights into microbial ecological concepts

74 including novel niches and community assembly[13]. Experimentally, the microbial communities were

75 monitored using 16S-ribosomal RNA (rRNA) and messenger RNA (mRNA) metatranscriptomic non-

76 target sequencing. Recently, 16S-rRNA sequencing has become an established method for analyzing

77 bacterial communities[25,26] in biotechnological applications, with detailed sample preparation and data-

78 processing pipelines available[27]. To complement the taxonomic survey, mRNA sequencing (RNAseq) was

79 performed to determine the functional profiles of the communities, for which Enzyme Commission (EC)

80 numbers were used as a proxy for an expressed function. Numerically, Monod growth kinetics were

81 employed in a simplistic MRT-diversity model to provide a concrete mechanistic basis for the connection

82 between the MRT and the community composition within the SBR. This model uniquely investigates the

83 underlying available range of growth parameters that result in persistence within the community.

84 Organisms must survive through substrate-rich and –poor conditions within the SBR, suggesting the main

85 tradeoff in survival strategies available to the organism is the ability to capture resources versus the ability

86 to withstand starvation. Therefore, individual distinctive combinations of the maximum growth rate (μ_{max})

87 and endogenous decay rate (b_e) are modelled and considered to be bounded in an ecological range of

88 permissible values. Critically, allowing b_e, a parameter that has been previously shown to be species-

89 specific[28], to vary between community members allows for the coexistence of multiple μ_{max} values within

90 a given community. This novel approach permits observing how the MRT, the independent factor in the

91 model, influences the available set of μ_{max} and b_e values. This simplified view of community composition

92 leads to a better conceptual understanding of the influence of the MRT on the richness and diversity of

93 microbial communities in the studied system.

Materials and Methods

Activated Sludge Reactor Configuration

Briefly, six automated sequencing batch reactors (6 × 12 L) treating local municipal wastewater after primary clarification were operated in parallel at MRTs of 1, 3, 5, 7, 10, or 15 days (d) as detailed previously[29] and summarized in Supplemental Table 1. Forty-eight days (time-point 1; TP1) and 187 days (time-point 2, TP2) after start-up, activated sludge samples were collected for DNA (at the start of the previously described biotransformation experiment[29]) and RNA (5 hours after the start of the experiment) extraction.

Sample Collection

To collect samples, culture (two 20-mL samples for TP1 and a 20-mL and 40-mL sample for TP2 for DNA and RNA analysis, respectively) was withdrawn and centrifuged at 3,345 x g for 10 min at 4ºC. The supernatant was then discarded and the pellets were stored at -80ºC until further processing.

DNA/RNA Isolation

The total RNA and genomic DNA isolation protocol consisted of a phenol:chloroform:isoamylalcohol extraction followed by either a DNA PowerCleanup PRO Kit (Qiagen, Venlo, Netherlands) or MoBIO RNA Pro Clean-Up Kit (MoBio, Carlsbad, CA, USA) and purification with a TURBO DNase step (ThermoFisher Scientific, Waltham, MA, USA) as detailed in the Supplemental Materials and modified from Johnson et al. 2015. The RNA pellet was re-suspended with diethyl pyrocarbonate (DEPC) treated RNase-free water to a total volume of 50 μl. DNA samples were quantified on a Qubit (Invitrogen, Waltham, MA, USA) analyzer following the manufacturer's instructions, whereas RNA samples were quantified on a Nanodrop (Invitrogen) and quality-checked on a Bioanalyzer 2000 (RNA 6000 kit; Agilent Technologies, Santa Clara, CA, USA).

120 *16S Library Preparation and Sequencing*

121   In preparation for the 16S sequencing, The total RNA was reverse transcribed into complementary

122 DNA (cDNA) using the Superscript III Kit (Invitrogen) with random hexamer primers following the

123 manufacturer's instructions. The genomic DNA (gDNA) was used directly after the purification described

124 above.

125

126   The 16S-rRNA or -rDNA amplicon library preparation followed a standard procedure for the

127 Illumina MiSeq platform (Illumina, San Diego, CA) that is detailed in Supplemental Materials 1. Two sets

128 of 16S-rRNA primers (Integrated DNA Technologies, Inc., Skokie, Illinois, USA) were used in this

129 analysis to amplify the sample cDNA and gDNA to account for the potential for the bias of a single

130 primer-set[30]. The details of primers B1 and B2 are provided in Supplemental Table 2. The samples were

131 sequenced using the PE 300 method on a MiSeq platform (Illumina) at the Genomics Diversity Centre at

132 ETH Zurich, Switzerland. The raw data is publically available at EMBL-EBI under the study number

133 PRJEB22087. The read count per sample and associated rarefaction curves are presented in Supplemental

134 Figures 1 and 2, respectively.

135

136 *16S rRNA and rDNA Sequencing Data Processing and Analysis*

137   The raw data was checked for quality using FastQC[31] v0.11.2. The reported nucleic sequence of

138 the reads was then trimmed using PRINSEQ-lite[32] v0.20.4 to a length of 295 bp and merged using

139 USEARCH[33] v8.1.1756 (with a minimum overlap of 15 bp, minimum merge-length of 100, and a

140 maximum error of 5 bp). The primers were trimmed from the merged read using cutadapt[34] v1.5 with

141 wildcards allowed, a full-length overlap, and an error rate of 0.01. The reads were then filtered using

142 PRINSEQ-lite with an amplicon range of 431-506 and 252-254 for B1 and B2, respectively, a minimum

143 quality mean of 15, and no ambiguous nucleotides allowed. USEARCH was employed to denoise the

144 reads into exact sequence variants (ESVs; zero-level operational taxonomical units, ZOTUs using

145 UNOISE3) and assign taxonomic origin (using usearch_global, 70% identity against the SILVA 16S

146   database (release 128), followed by sintax with a 70% identity cutoff). The total number of raw and

147   cleaned reads per sample for the B2 primer ranged from 59,311-191,897 with a median of 113,347 and

148   54,864-185,331 with a median of 104,537, respectively (the details for every sample are provided for

149   primers B1 and B2 in Supplemental Tables 3 and 4, respectively). In total, 99.3% of these reads were

150   binned into 10,644 ESVs, with 2,918 ESVs displaying more than 10 reads in at least one sample. Primer

151   B2 is considered in the main text because more positive and negative controls were analyzed for B2 than

152   B1.

153

154   The resulting data was then analyzed in R v3.5.1 using phyloseq[35] v1.24.2 as detailed in

155   Supplemental File 1. The bacterial 16S richness (^{0}D; on rarefied data to remove potential sampling effort

156   effects) and Shannon diversity index ($\ln(^1$D$)$; on non-rarefied data) were calculated as n and

157    $\exp(-\sum_{i=1}^{n} p_i * \ln(p_i))$, respectively, where n is the number of ESVs and p_i is the abundance-weighted

158   proportion of ESV i [36-38]. When relating metrics throughout this study, Spearman rank-correlation

159   (denoted as r) analyses were employed to avoid imposing assumptions of linearity.

160

161   *RNAseq Library Preparation and Sequencing*

162   The RNA samples were processed into libraries and sequenced following the Illumina TruSeq

163   Single-End-Read 150 bp pipeline of the Genomics Facility at the University of Basel. In brief, the

164   abundant ribosomal sequences in the samples were degraded to enhance the mRNA fraction using the

165   Ribo-Zero Gold Epidemiology Kit (Illumina) to target Eukaryotic, Bacterial, and Archaeal sequences.

166   During testing, this Epidemiology Kit was found to outperform a sequential application of the Ribo-Zero

167   Gold Bacterial and Eukaryotic Kits (Illumina) on the activated sludge samples (81.3±5.2% versus

168   15.4±1.8% of resulting reads of non-rRNA origin). The adapter addition, sample cleanup, and fragment

169   selection were performed as outlined in the Illumina TruSeq protocol. The samples were then sequenced

170   on a NextSeq 500 Platform (Illumina pipeline 2.4.11). The raw data are publically available at EMBL-EBI

171 under the study number PRJEB22087. The quality of the RNA as extracted, RNA after depletion, and

172 resulting fragments are provided in Supplemental Figures 3-5.

173

174 *RNAseq Data Processing, Normalization, and Analysis*

175   The raw read files were trimmed of adapter sequences, index sequences, and low quality reads

176 using Trimmomatic[39] v0.33. The raw and trimmed reads were also checked for quality using FastQC[31]

177 (Supplemental Figure 6-9). To remove contaminating rRNA reads *in silico*, the trimmed reads were

178 compared against rRNA databases (Silva version 119 (Bacteria 16S & 23S, Archaea 16S & 23S,

179 Eukaryote 18S & 28S) and RFAM (5S & 5.8S)) and filtered using SortMeRNA[40]. Sequences passing the

180 quality control were annotated with the descriptors provided in the Enzyme Commission (EC) Number

181 Uniprot database using DIAMOND[41] v0.2.1 with the blastx command and a minimum bitscore cutoff of

182 50 (all other parameters set to their default). Because we are primarily investigating EC annotation that

183 can be shared across taxa and not specific genes from individual species, only the best annotation per read

184 was recorded. The full Uniprot-TrEMBL database was created by downloading the database on March 6th,

185 2018 (36.8 billion amino acids in 109 million sequences). The narrower Uniprot-EC database was created

186 by searching for ec:* and downloading all matching hits on March 6th, 2018 (5.8 billion amino acids in

187 14.9 million sequences). The script required to process the raw RNAseq files, generate the database,

188 annotate the reads, and extract the taxonomic Uniprot identifiers is provided as Supplemental File 2. The

189 resulting raw sequencing files contained 41.8-54.4 million reads, of which 72.5-87.7% remained in the

190 dataset after quality and rRNA filtering. In total, 32.8-47.7 million reads per sample were submitted for

191 annotation, resulting in 5.1-9.8 million reads being annotated per sample (Supplemental Table 5;

192 Supplemental Figures 10-11).

193

194   The read counts were aggregated per EC number, and these EC numbers were used as a proxy of

195 the functional profile in this study. When the Uniprot entry that provided the annotation of a read

196 maintained multiple EC numbers, the read was assigned equally to each EC number (<5% of all

197 annotations maintained multiple EC designations). The rarefaction curves showed that the richness of EC

198 numbers saturated within the library's sequencing depth (Supplemental Figure 12). For normalization, the

199 count data was treated compositionally in that the abundance of a specific EC number was divided by the

200 total number of reads identified to encode a protein. The total number of protein encoding reads was

201 determined by first using 500 k reads from each library to search against the full Uniprot-TrEMBL

202 database and then multiplying the fraction annotated with the total number of reads submitted to the

203 Uniprot-EC database (Supplemental Table 5; Supplemental Figure 13).

204

205 ***MRT-Diversity Model Construction***

206 *i.   MRT-Diversity Model Approach, Assumptions, and Limitations*

207 In the MRT-diversity model, Monod-type bacterial growth mathematics[42] were employed

208 dynamically to approximate the linkage between the MRT and community composition in the

209 experimental reactor (Figure 1.a). Monod-kinetics use the μ_{max}, b_e, substrate affinity (K_s), and yield (Y) to

210 describe the growth of an organism's biomass (X) on a given substrate (S). The approach presented here

211 utilizes these parameters in a novel manner by exploring the range of their combined values (an

212 approximation of community diversity) that leads to persistence over a MRT gradient.

213

214 In our approach, we apply a number of simplifications to typical considerations employed in other

215 Monod-growth based dynamic models[24] to determine the range of growth parameters leading to

216 persistence. Specifically, the wastewater is considered a single substrate (e.g., no distinction of

217 carbonaceous or nitrogenous compounds), growth limitations resulting from sources other than substrate

218 availability are considered constant (e.g., mass transfer, toxic product formation, additional substrates),

219 and competition is allowed only for this single substrate. When triggered, assigned flow rates and influent

220 composition are also assumed to be temporally stable to remove variability resulting from other

221 independent variables, and mixing within the reactors is considered perfect (except during the settle

222 phase). Changes in steady-state growth depend only on the maximal gene expression and enzyme kinetics,

223   and the availability of the enzyme pool is considered temporally stable thereby neglecting evolution. In

224   turn, this stability is assumed to allow instantaneous adjustments of the growth rate to the change in the

225   substrate concentrations (i.e., time lags have elapsed).

226

227   To capture competition over both substrate-rich and –poor phases, individuals within this model

228   are allowed to be distinct in two growth parameters only: μ_{max} and b_e. Both internal (consumption of stored

229   substrate) and external (adverse environmental conditions, cell programmed death, and viral attack) decay

230   are considered incorporated in the b_e parameter[43]; higher order ecological considerations dependent on the

231   consideration of additional substrates such as predation and growth on lysis products are excluded from

232   the model. The maximum and minimum μ_{max} and b_e values are bounded by ecological limits, and a

233   constraining combination of growth parameters must be satisfied by the range. In summary, the main

234   uncertainties in the model include the appropriateness of restricting the analysis to a single substrate, the

235   placement of the constraining growth parameters, the uncertain assignment of the ecological constraints,

236   and the exclusion of other contributors to diversity such as population oscillation and the time to reach

237   equilibrium[44].

238

239   *ii.   Role of μ_{max} and b_e in the MRT-Diversity Model*

240   To conceptualize the interaction of the μ_{max}, b_e, and MRT within the MRT-Diversity model of a

241   SBR, the solution for the minimum substrate concentration (S^{*}_{min}) that leads to persistence in a

242   continuously stirred tank reactor[45] (Supplemental Materials 2) provides a simplified analogy that can be

243   written without including differential equations:

244    $$S^{*}_{min} = \frac{(1 + b_e * MRT)}{(\mu_{max} - b_e) * MRT - 1} K_s \qquad \text{Eq. 1}$$

245   where all parameters were defined previously. Organisms with the lowest calculated S^{*}_{min} values will

246   persist in the reactor because they will outcompete other community members for the sole resource. In

247   previous models, a single surviving species would be selected because of the hypothesized inability of

248   other organisms to exactly match the μ_{max}, K_s, and b_e combination required for persistence in the reactor[45].

249   Notably, we relax this constraint and allow multiple organisms to grow on a single substrate. Modelling

250   co-existing combinations of growth parameters explores whether we can predict richness and diversity

251   values similar to the experimentally observed values over a MRT gradient.

252

253   In developing this Monod-kinetics model of multiple organisms for the investigated SBR (Figure

254   1.a), the combination of growth parameters that are allowed at a given MRT is simply given by the

255   maximum and minimum μ_{max} and b_e values that persist in the reactor to represent competition in both the

256   resource-rich and –poor operating phases (Figure 1.b). The K_s influences the μ_{max} values leading to

257   persistence in the reactor less than b_e (Supplemental Figure 14) because it represents another component

258   of resource capture rather than variations in survival during low or no production. K_s was therefore held

259   constant to minimize model complexity. The line of growth parameter combinations that results in

260   persistence (and determined by the equations detailed below) is required to fall within a roughly set

261   ecological range and to pass through constraining growth parameters ($\mu_{max,constrain}$ and $b_{e,constrain}$; values that

262   are initially assumed to remain unchanged between reactors, arbitrarily set to the center of the range, and

263   explored further in Supplemental Figure 15). To establish the permissible ecological values of b_e, the

264   extremes of previously reported observations (from ~0.02 [47,48] to ~0.2 d^{-1} [49]) were used as approximate

265   boundaries ($b_{e,eco}$; 0.02 to 0.2 d^{-1}), and an average value (0.11 d^{-1}) was selected as the $b_{e,contrain}$ (Table 1).

266   The $\mu_{max,eco}$ boundaries (0.2 to 9.8 d^{-1}) were set to exceed the range of values reported for a previous MRT

267   gradient[50], and an average value of 5 d^{-1} was selected as the $\mu_{max,contrain}$ (Table 1). The model was found to

268   be rather insensitive to the selection of these constraining points (Supplemental Figure 15).

269

270   *iii.   SBR Differential Equations*

271   The combination of growth parameter values resulting in persistence (i.e., non-zero steady-state

272   concentrations) across the ecological range were determined with the following system of differential

273   equations that describes the flow, biomass, and substrate concentrations within the SBR (Figure 1.a):

274    $\frac{dV}{dt} = Q_{in} - Q_{clarified\ drain} - Q_{mixed\ drain}$ Eq. 2

275 $$\frac{dX_i}{dt} = \left(-X_i(t) * Q_{mixed\ drain}/V(t) + \left[\frac{\mu_{max,i}*S(t)}{K_s+S(t)} - b_{e,i}\right] * X_i(t)\right)\Big|_{i=1}^{n}$$ Eq. 3

276 $$\frac{dS}{dt} = \left(S_{in} * \frac{Q_{in}}{V(t)} - (Q_{clarified\ drain} + Q_{mixed\ drain}) * S(t)/V(t) - \sum_{i=1}^{n}\left[\frac{\mu_{max,i}*S(t)}{K_s+S(t)}\right] * \frac{X_i(t)}{Y}\right)$$ Eq. 4

277 where the flowrates are triggered during their respective cycles (and are zero otherwise); the *i* subscript

278 indicates parameters and biomass for the ith combination of growth parameters (ranging from 1 to n) that

279 were modelled simultaneously; and all other parameters are defined in Table 1 and further described in

280 Supplemental File 3. To ensure flow balance across the SBR cycle (Figure 1.a), the $Q_{clarified\ drain}$ is

281 calculated to offset the Q_{in} and $Q_{mixed\ drain}$ (outflow of suspended biomass):

282 $$Q_{clarified\ drain} = \frac{Q_{in}*t_{in} - Q_{mixed\ drain}*t_{mixed\ drain}}{t_{clarified\ drain}}$$ Eq. 5

283 where all parameters are defined in Table 1. Notably, the MRT is determined as the full volume of the

284 reactor divided by the total volume of suspended biomass removed ($Q_{mixed\ drain}*t_{mixed\ drain}$) per six cycles

285 (one day).

286

287   An iterative approach was used to calculate the μ_{max} values resulting in persistence for nine b_e

288 values distributed across the ecological range. A full solution line was then fit to these nine points (Figure

289 1.b). This solution line was found to depend only on those parameters directly influencing μ_{max}, b_e, and

290 MRT and was insensitive to changes in other global parameters such as the S_{in}, Y, and K_s. All differential

291 equations mentioned in this study were analyzed using deSolve[51] v1.21, and all calculations were

292 performed in R v3.5.1 (Supplemental File 3).

293

294   *iv.*   *MRT-Diversity Model Alpha Diversity Calculation*

295   After determining the solution (Figure 1.b), the length of the line representing all combinations of

296 μ_{max} and b_e leading to survival within the reactor was then calculated:

297 $$\text{Growth Parameter Solution Length} = \sqrt{\left(\frac{\mu_{max,max}-\mu_{max,min}}{\mu_{max,eco,max}-\mu_{max,eco,min}}\right)^2 + \left(\frac{b_{e,max}-b_{e,min}}{b_{e,eco,max}-b_{e,eco,min}} * \text{Scaling Factor}\right)^2}$$ Eq. 6

298 where the Scaling Factor is set to 0.25 to represent a case when the μ_{max} range contributes more to the

299 length than the b_e (emphasizing the fact that b_e serves more to allow the coexistence of different μ_{max}

300   values rather than contribute to diversity directly; see Supplemental File 3). The growth parameter

301   solution length is utilized as a proxy for the richness of a community; this length will most likely be an

302   underestimate of true richness as a result of binning organisms (or ESVs when comparing to 16S data) that

303   display the identical combination of growth parameters. The Shannon diversity index was determined by

304   numerically solving differential equations for the steady-state biomasses (X_i(steady-state)) when

305   considering the number of distinct combinations of growth parameters within the community to be the

306   length of the range multiplied by a constant value (n = 50; discretionarily set to achieve an integer value

307   representative of community size and a timely computation of the differential equations). The

308   instantaneous substrate utilization rate (k_{theo}) was calculated as the maximum substrate utilization rate

309   determined at the beginning of one cycle.

Results and Discussion:

Observed Taxonomic Richness and Diversity Increases with MRT

The ESV richness increases monotonically across the MRT gradient for the active community members, i.e., the 16S rRNA (Spearman rank correlation r = 0.98 and 0.89 for TP1 and TP2, respectively), but displays a lower correlation to a monotonic trend for the present community members, i.e., the 16S rDNA (r = 0.81 and 0.77 for TP1 and TP2, respectively)(Figure 2.a & b). Additionally, the abundance weighted diversity metric, the Shannon diversity index, shows a decelerating increase in the rRNA transcripts which levels off above 5.1. Overall, the observed increase in the Shannon diversity value between 3 to 10 d (i.e., a mean±s.d. of 4.6±0.27 to 5.3±0.20, respectively) agrees with a previous study investigating lab-scale synthetic wastewater-treating membrane bioreactors (MBRs)[52]. Other studies utilizing synthetic wastewater indicated no substantial difference between the community diversity metrics at ~2 and 10 d MRT in a MBR system[6,7], or a decrease in diversity from 3 to 8 d MRT in a SBR system[53]. These studies also employed other sequencing techniques such as denaturing gel gradient electrophoresis[7] and terminal restriction length polymorphism analysis[53] that can affect the exact quantified values, but are not expected to affect the reported trends of stable or decreasing diversity metrics. By contrast, investigations of full-scale WWTPs reported a comparable increase in diversity metrics at longer MRT[4,5,54], suggesting that real wastewater is required to consistently display a direct MRT-diversity relationship as also observed here.

Of the fifteen highlighted orders (Figure 2.c & d), six (*Burkholderiales*, *Rhodocyclales*, *Myxococcales*, *Sphingobacteriales*, *Rhodobacteriales*, and *Pseudomonadales*) were previously demonstrated to be commonly shared by a wide variety of activated sludge[55,56]. Across a set of 13 Danish WWTPs, genera of the *Thiotrichales* order were abundantly observed in only two WWTPs, highlighting the potential transient nature of this population in WWTPs[56]. At both time-points, the relative 16S rRNA transcript abundance of *Burkholderiales* decreases by nearly a factor of two with increasing MRT (from 41.3±0.30% to 19.6±0.23% and 37.5±0.60% to 21.1±0.15% of the community for TP1 and TP2,

336 respectively) consistent with a previous study[4], whereas *Rhodocyclales* (from 21.2±0.44% to 41.3±0.20%

337 and 12.3±0.21% to 19.6±0.59%) and *Myxococcales* (from 0.11±0.30% to 12.2±0.38% and 0.14±0.01% to

338 5.7±0.12%) show increasing abundances. Additionally, a low abundance subpopulation capable of

339 oxidizing ammonia to nitrate, the *Nitrosomonadales*, established at longer MRT when nitrification was

340 noted[28] and expected[57].

341

342   The relative distribution of the orders are maintained in both the TP2 16S rRNA and rDNA

343 profiles (Spearman r for the top 50 orders of 0.90, 0.88, 0.80, 0.74, 0.72, and 0.60 for 1, 3, 5, 7, 10, 15 d

344 MRT, respectively). However, the profile in TP1 was substantially more variable (r = 0.27, 0.33, 0.47,

345 0.47, 0.38, and 0.25, respectively). This divergence is attributed to the detection of unique orders

346 (*Caldilineales*, *Lactobacillales*, *Micrococcales*) and to the over-abundance of members within the

347 *Thiotrichales* order in the TP1 rDNA profile (Figure 2.c). This over-abundance suggests that the

348 filamentous *Thiotrichales* in TP1 causes a negative selection event in those reactors in which the most

349 dominant organism by biomass (rDNA) is not the most active or productive (rRNA)[58]. This more variable

350 signal also results in the *Nitrosomonadales* order displaying a 10- to 100-fold lower rDNA than rRNA

351 signal (Figure 2.c and d), obscuring the ability to detect the known MRT-dependent emergence of this

352 organism and its causal relationship to nitrification[59]. Overall, the 16S results suggest that the expression

353 of 16S rRNA is more reflective of activity than the detection of an organism (16S rDNA) in the activated

354 sludge experiments, supporting previous findings in surveys of other aerobic systems[60].

355

356 ***MRT is a Driver of Modeled Taxonomic Richness and Diversity***

357   In the constructed MRT-diversity model, increasing the MRT expands the range of combinations

358 of μ_{max} and b_e values that lead to persistence (Figure 1.b; Supplemental Table 6). Strikingly, the increase in

359 the range of growth parameters (Figure 3.a) strongly correlates with the 16S rRNA observed richness

360 (Figure 2.a and b) with r-values of 0.98 and 0.89 for TP1 and TP2, respectively (Figure 3.a inset). In terms

361 of previously developed theories of microbial ecology, reducing the community complexity into tradeoffs

362 between parameters that describe resource capture and represent survival during low-production (μ_{max} and

363 b_e, respectively; components of the Competition-Stress-Ruderal continuum[61,62]) recaptures the trend of

364 increasing richness across the MRT range for the SBR. When utilizing a variable parameter that is

365 independent of the substrate concentration (b_e), the maximum effect resulting from the difference in ability

366 to withstand low-production occurs when no substrate is present, a critical component to capture in the

367 dynamic model of SBRs. By comparison, when considering only variation in the parameters describing

368 resource capture (μ_{max} and K_s; components of the r/k-specialist model[63]; Supplemental Figure 14), the

369 influence of both μ_{max} and K_s diminishes as the substrate concentration approaches zero, challenging the

370 applicability of the resource-capture-only framework in SBR systems with extended react phases.

371

372   Although a range of growth parameters will persist (Figure 1.b), the organisms they represent will

373 be present at various abundances at steady-state (X_i in Figure 1.c). The Shannon diversity for the SBR

374 shows a decelerating increase with MRT (Figure 3.b), matching the 16S rRNA observed data ($r = 0.97$ and

375 0.74 for TP1 and TP2, respectively (Figure 3.b inset)). An underlying assumption in this comparison is

376 that the ratio of the biomass resulting from a given combination of growth parameters to the number of

377 representative rRNA transcripts is constant; however, this ratio varies even at the gene copy per genome

378 level[64]. Therefore, the general trend of the curve is informative of whether the MRT influences the

379 diversity, but the magnitude of the shifts would be substantially affected by this rRNA-to-biomass ratio.

380

381   The substrate consumption rate of the entire activated sludge community is often monitored

382 through respirogram bulk tests (i.e., biomass normalized maximal oxygen uptake rate (OUR) analyses)[65]

383 and has been previously reported to slow with increasing MRT[50,66], suggesting an adaptation of the

384 community. Our model allows predicting the instantaneous substrate utilization rate (k_{theo}), and the

385 previously published slowing trend is not observed from the initial default parameters (modelled k_{theo} of

386 5.0 and 5.8 d^{-1} at 1 and 15 d, respectively). This disconnect likely stems from underlying assumptions of

387 our model, most notably fixing constraining central growth parameters ($\mu_{max,constrain}$ and $b_{e,constrain}$).

388　　However, when using $b_{e,ecomax}$ as $b_{e,constrain}$ instead, the decreasing trend in the previously reported empirical

389　　values is successfully mirrored (k_{theo} of 4.8 and 4.1 d^{-1} at 1 and 15 d, respectively) while the diversity

390　　profiles are conserved. Therefore, ambiguity remains regarding the accurate placement of these

391　　controlling parameters as well as the ecological range parameter values.

392

393　　　　Overall, these observed and modelled results complement a previous study monitoring an

394　　activated sludge reactor for 313 days in a 30, 12, 3, 30 d MRT disturbance cycle, in which an increased

395　　diversity was noted for the higher MRT values[4,67]. In that study, two mechanisms were proposed that

396　　contribute to the higher richness and diversity: decreases in unconsumed resources (analogous to

397　　decreases in the S^{*}_{min} in Eq. 1 or an increase in the length of the starvation phase in the SBR) and increases

398　　in niche space (represented by the range of growth parameters and assumed to represent the richness in the

399　　MRT-diversity model). Our model uses the MRT and growth parameters to represent these two

400　　mechanisms in separate equations and predicts that the richness increases as the availability of

401　　unconsumed resources decreases across the observed MRT range (Figure 3). Notably, the MRT-richness

402　　profile can display a non-monotonic trend at higher MRT or by placing the constraining growth

403　　parameters ($\mu_{max,constrain}$ or $b_{e,constrain}$) close to the ecological limits. By contrast, the availability of

404　　unconsumed resources will consistently decrease with increasing MRT. Notably, observing a non-

405　　monotonic MRT-richness relationship would suggest that niche space contributes more to diversity. In

406　　future studies, the potential for this non-monotonic profile should be tested by establishing reactors

407　　exceeding the maximum MRT observed here (Figure 2.a and b). Additionally, uncovering a transition

408　　from a monotonic to non-monotonic MRT-richness relationship would assist in testing the underlying

409　　assumptions of the model and more accurately estimating the ecological ranges and constraining growth

410　　parameters.

411

412　　*Observed Functional Richness, but not Diversity, Increases with MRT*

413   With increasing MRT, conceptually either an organism absent at lower MRT may occupy the

414   additional growth parameter space or a shared microorganism across MRT expresses different functional

415   enzymes. To test for shifts in the functional profile, the metatranscriptomes of the experimental

416   communities were sequenced, annotated as EC numbers, and analyzed using alpha diversity indicators of

417   the number of unique (richness) and evenness of the relative abundance (Shannon diversity index) of the

418   EC sub-subclasses and numbers (Figure 4). Similar to what was noted in a previous study, the 16S rRNA

419   taxonomic and EC number richness of both time-points display a strong correlation ($r = 0.98$ and 0.97 for

420   the TP1 and TP2 samples, respectively)[12] and a nearly monotonic increase with MRT in richness (Figure

421   2.a and b; Figure 4). This relationship between the taxonomic and functional richness is not as consistently

422   strong with the 16S rDNA ($r = 0.81$ and 0.94, respectively). These results again highlight that a measure

423   more reflective of current activity within the cell (16S rRNA) links better with the overall functional

424   profile (mRNA) than a survey of presence alone (16S rDNA). When considering the relative abundances

425   of a given EC sub-subclass or number, the Shannon diversity displays a non-monotonic profile,

426   contrasting the taxonomic profiles and the EC sub-subclass or number richness. This disconnect between

427   taxonomic and functional diversity has been previously demonstrated in model wastewater reactors[13],

428   suggesting that this result represents a true signal beyond simple limitations with the 16S rRNA measure

429   (e.g., abundance not always correlating with growth rate, inter-species differences in copy numbers per

430   cell[68], steep ecological gradients across the SBR cycle). However, the disconnect between the functional

431   richness and Shannon diversity indicates that although the quantity of EC sub-subclasses or numbers

432   increases across the gradient, specific categories increase in dominance at longer MRTs, offsetting the

433   increased richness (Figure 4).

434

435   ***The Relative Abundance Shifts of "Rare" Enzyme Classes Drive the Functional Diversity Profile***

436   Several overrepresented EC sub-subclasses in terms of observed abundance (Figure 5.a), e.g., the

437   2.7.7 nucleotidyltransferase (containing 2.7.7.6 RNA-polymerase, RpoB) and 5.99.1 other-isomerase

438   (containing 5.99.1.2 DNA topoisomerase) EC sub-subclass, decrease in their fractional share of the

439   metatranscriptome as the MRT increases (Figure 5.b). Simultaneously, the fractional share of sub-

440   subclasses that include oxidoreductases and nitrogen-processing related enzymes that are linked to the

441   emergence of nitrification over the MRT gradient (e.g., nitrogenous oxidoreductase with [1.7.2] and

442   without [1.7.99] cytochrome) increase (Figure 5.a and b). The over-abundance of nitrogen metabolism-

443   related gene transcripts has been previously noted in activated sludge even when nitrifiers are a minor

444   fraction of the community[69]. Notably, these EC sub-subclasses and associated numbers that markedly

445   increase in abundance over the MRT induced the non-monotonic functional diversity profile, indicating

446   that substantially different abundances of mRNAs encoding for specific enzymes are likely required to

447   achieve those growth parameters resulting in persistence. The discrepancy between the non-monotonic

448   functional diversity profile of all EC numbers and the monotonic taxonomic diversity profile for a single

449   targeted transcript (e.g., 16S-rRNA) results from enzymes displaying additional properties[70] such as

450   specific substrates affinities, product turnover rates[71], and protein-to-transcript ratios[72,73] that affect their

451   relative fractional abundances.

452

453   To compare the observed diversities within each EC sub-subclass (similar to the 16S-rRNA

454   analysis), specific taxonomic richness and diversity values were calculated based on the putative genus-

455   level organism-origin annotation that each mRNA read is assigned (Figure 5.c and d, respectively).

456   Focusing on the aforementioned 2.7.7 and 1.7.99 sub-subclasses to highlight categories demonstrated to

457   be common and rare, respectively (Figure 5.c), the diversity profile of the common EC sub-subclass 2.7.7

458   displays a positive relationship with that of the 16S rRNA (Figure 2; r = 0.89 and 0.59 for TP1 and TP2

459   Shannon indices, respectively). In contrast, a divergent profile is seen for the taxonomic diversity of the

460   rare 1.7.99 sub-subclass (r = -0.65 and -0.76 for TP1 and TP2, respectively), indicating that select

461   organisms dominate the origin of the reads within this category at higher MRTs. Notably, reads from the

462   nitrifying *Nitrospira*[74] dominate the 1.7.99 sub-subclass. The nitrification rate intensifies across the MRT

463   gradient, suggesting that *Nitrospira* expressed the proper bulk-growth parameters to persist and thrive

464   within the community. The greater share of the overall reads transcribed resulting from a single,

465 nitrification-related organism contributed to the noted decrease in functional diversity of the overall

466 community (Figure 4) .

467

468 When further binning reads into Domain-level taxonomic origin, a substantial fraction of

469 annotations originating from Eukaryotic organisms (Figure 5.e) were noted for certain EC sub-subclasses

470 increasing in abundance over the MRT gradient (Figure 5.a). In activated sludge, increasing MRT over the

471 studied range have been reported to promote a higher abundance of Protozoa[75], organisms that are

472 overlooked in bacterial-targeted taxonomic surveys of WWTPs. This signal in the mRNA data could

473 confound the previous comparison between the taxonomic and functional diversity metrics. However,

474 when reanalyzing the functional diversity metrics (Figure 4) for the Bacterial portion (Eukaryotic filtered),

475 a similar profile is obtained (Supplemental Figure 19), supporting the detection of a true distinction

476 between the taxonomic composition and the functional profile.

477

478 **Conclusions**

479 As demonstrated experimentally, increasing the MRT positively affects the taxonomic richness

480 and diversity as well as functional richness of the monitored activated sludge community. To

481 conceptualize these findings, a naïve model was constructed that utilized Monod-kinetics in a novel

482 manner by considering wastewater as a single substrate and the community as a collection of growth

483 parameter combinations. Combinations of μ_{max} and b_e values were selected to represent the tradeoff

484 between resource capture and survival during low production, reflecting considerations within the C-S-R

485 description of microbial ecology. This MRT-diversity model predicted that the range of μ_{max} and b_e values

486 expands with increasing MRT for the studied system, suggesting a new, kinetic parameter-driven metric

487 that correlates strikingly well with the observed taxonomic profile and the functional richness across the

488 MRT gradient. For a new community member to occupy these opened growth parameter combinations

489 and thereby increase the taxonomic richness, previously unobserved EC numbers are likely required

490 because of the noted increases in functional richness. In contrast to the taxonomic abundance-weighted

491 diversity, the functional diversity displayed a non-monotonic trend over the MRT range. Whereas more

492 EC sub-subclasses and numbers are detected at higher MRTs, their fractional share of the overall activity

493 of the community varies depending on the expressed function. For example, rare sub-subclasses related to

494 nitrification substantially increase in dominance at longer MRTs in this system. Although the complexity

495 of the relationship between EC numbers is not successfully captured, the simplification of the community

496 into combinations of μ_{max} and b_e values appears to be a useful approximation for predicting changes in

497 taxonomic richness and diversity as well as functional richness over a MRT gradient in this system.

498 Because this study is the first to employ Monod-kinetics in this manner, future work should determine

499 whether the approach and assumptions introduced here are valid when used to describe other systems,

500 explore the concepts of the constraining growth-parameter combination and ecological boundary values,

501 and subdivide influent resources into individual substrate types (e.g., nitrogen-containing compounds).

502

503 **Conflict of Interest:**

504 The authors declare no conflicts of interest.

505

513

References:

1. Curtis, T. P., Sloan, W. T., and Scannell, J. W. (2002). Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences*, 99(16), 10494-10499.

2. Locey, K. J., and Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences*, 113(21), 5970-5975.

3. Antwis, R. E., Griffiths, S. M., Harrison, X. A., Aranega-Bou, P., Arce, A., Bettridge, A. S., et al. (2017). Fifty important research questions in microbial ecology. *FEMS Microbiology Ecology*, 93(5).

4. Vuono, D. C., Benecke, J., Henkel, J., Navidi, W. C., Cath, T. Y., Munakata-Marr, J., et al. (2015). Disturbance and temporal partitioning of the activated sludge metacommunity. *The ISME Journal*, 9(2), 425-435.

5. Meerburg, F. A., Vlaeminck, S. E., Roume, H., Seuntjens, D., Pieper, D. H., Jauregui, R., et al. (2016). High-rate activated sludge communities have a distinctly different structure compared to low-rate sludge communities, and are less sensitive towards environmental and operational variables. *Water Research*, 100, 137-145.

6. Bagchi, S., Tellez, B. G., Rao, H. A., Lamendella, R., and Saikaly, P. E. (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(5), 2361-2370.

7. Akarsubasi, A. T., Eyice, O., Miskin, I., Head, I. M., and Curtis, T. P. (2009). Effect of sludge age on the bacterial diversity of bench scale sequencing batch reactors. *Environmental Science & Technology*, 43(8), 2950-2956.

8. Fang, H., Chen, Y., Huang, L., and He, G. (2017). Analysis of biofilm bacterial communities under different shear stresses using size-fractionated sediment. *Scientific Reports*, 7(1), 1299.

9. Roager, H. M., Hansen, L. B., Bahl, M. I., Frandsen, H. L., Carvalho, V., Gøbel, R. J., et al. (2016). Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nature Microbiology*, 1(9), 16093.

10. Kreuzinger, N., Clara, M., Strenn, B., and Kroiss, H. (2004). Relevance of the sludge retention time (SRT) as design criteria for wastewater treatment plants for the removal of endocrine disruptors and pharmaceuticals from wastewater. *Water Science and Technology*, 50(5), 149-156.

11. Falås, P., Andersen, H. R., Ledin, A., and la Cour Jansen, J. (2012). Impact of solid retention time and nitrification capacity on the ability of activated sludge to remove pharmaceuticals. *Environmental Technology*, 33(8), 865-872.

12. Johnson, D. R., Lee, T. K., Park, J., Fenner, K., and Helbling, D. E. (2015). The functional and taxonomic richness of wastewater treatment plant microbial communities are associated with each other and with ambient nitrogen and carbon availability. *Environmental Microbiology*, 17(12), 4851-4860.

13. Pholchan, M. K., Baptista, J. D. C., Davenport, R. J., Sloan, W. T., and Curtis, T. P. (2013). Microbial community assembly, theory and rare functions. *Frontiers in Microbiology*, 4.

14. Louca, S., Parfrey, L. W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science*, 353(6305), 1272-1277.

15. Louca, S., Jacques, S. M., Pires, A. P., Leal, J. S., Srivastava, D. S., Parfrey, L. W., et al. (2017). High taxonomic variability despite stable functional structure across microbial communities. *Nature Ecology & Evolution*, 1(1), 0015.

16. Frossard, A., Gerull, L., Mutz, M., and Gessner, M. O. (2012). Disconnect of microbial structure and function: enzyme activities and bacterial communities in nascent stream corridors. *The ISME Journal*, 6(3), 680-691.

17. Purahong, W., Schloter, M., Pecyna, M. J., Kapturska, D., Däumlich, V., Mital, S., et al. (2014). Uncoupling of microbial community structure and function in decomposing litter across beech forest ecosystems in Central Europe. *Scientific Reports*, 4, 7014.

18. Kivlin, S. N., and Hawkes, C. V. (2016). Temporal and spatial variation of soil bacteria richness, composition, and function in a neotropical rainforest. *PloS One*, 11(7), e0159131.

19. Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences*, 108(34), 14288-14293.

20. Boon, E., Meehan, C. J., Whidden, C., Wong, D. H. J., Langille, M. G., and Beiko, R. G. (2014). Interactions in the microbiome: communities of organisms and communities of genes. *FEMS Microbiology Reviews*, 38(1), 90-118.

21. Daims, H., Taylor, M. W., and Wagner, M. (2006). Wastewater treatment: a model system for microbial ecology. *Trends in Biotechnology*, 24(11), 483-489.

22. Downing, A. L., and Hopwood, A. P. (1964). Some observations on the kinetics of nitrifying activated-sludge plants. *Schweizerische Zeitschrift für Hydrologie*, 26(2), 271-288.

23. Lawrence, A. W., and McCarty, P. L. (1970). Unified basis for biological treatment design and operation. *Journal of the Sanitary Engineering Division*, 96(3), 757-778.

24. Henze, M., Gujer, W., Mino, T., and Van Loosdrecht, M. C. M. (2000). Activated sludge models ASM1, ASM2, ASM2d and ASM3. IWA publishing.

25. Degnan, P. H., and Ochman, H. (2012). Illumina-based analysis of microbial community diversity. *The ISME Journal*, 6(1), 183-194.

26. Sinclair, L., Osman, O.A., Bertilsson, S., and Eiler, A. (2015) Microbial community composition and diversity via 16S rRNA Gene Amplicons: Evaluating the Illumina Platform. *PLoS One* 10(2): e0116955.

27. Knight, R., Vrbanac, A., Taylor, B.C., Aksenov, A., Callewaert, C., Debelius, J., et al. (2018) Best practices for analysing microbiomes. *Nature Reviews Microbiology*, 16, 410–422.

28. van Bodegom, P. (2007). Microbial maintenance: a critical review on its quantification. *Microbial Ecology*, 53(4), 513-523.

576   29. Achermann, S., Falas, P., Joss, A., Mansfeldt, C., Men, Y., Vogler, B., et al. (2018) Trends in micropollutant

577   biotransformation along a solids retention time gradient. *Environmental Science & Technology,* 52(20),11601–11611

578   30. Guo, F., Ju, F., Cai, L., and Zhang, T. (2013). Taxonomic precision of different hypervariable regions of 16S rRNA

579   gene and annotation methods for functional bacterial groups in biological wastewater treatment. *PloS One*, 8(10),

580   e76185.

581   31. Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

582   32. Schmieder, R., & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*,

583   27(6), 863-864.

584   33. Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19), 2460-2461.

585   34. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet*, 17(1):10.

586   35. McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of

587   microbiome census data. *PloS One*, 8(4), e61217.

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

589   37. Chao, A., Chiu, C. H., and Jost, L. (2010). Phylogenetic diversity measures based on Hill numbers. *Philosophical

590   Transactions of the Royal Society B – Biological Sciences*, 365(1558), 3599-3609.

591   38. Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., et al. (2014). Rarefaction and

592   extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological

593   Monographs*, 84(1), 45-67.

594   39. Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data.

595   *Bioinformatics*, 30(15), 2114-2120.

596   40. Kopylova, E., Noé, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in

597   metatranscriptomic data. *Bioinformatics*, 28(24), 3211-3217.

598   41. Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature

599   Methods*, 12(1), 59-60.

600   42. Monod, J. (1949). The growth of bacterial cultures. *Annual Reviews in Microbiology*, 3(1), 371-394.

601   43. Liu, G., and Wang, J. (2015). Modeling effects of DO and SRT on activated sludge decay and production. *Water

602   Research*, 80, 169-178.

603   44. Saikaly, P. E., and Oerther, D. B. (2004). Bacterial competition in activated sludge: theoretical analysis of varying

604   solids retention times on diversity. *Microbial Ecology*, 48(2), 274-284.

605   45. Hsu, S. B., Hubbell, S., and Waltman, P. (1977). A mathematical theory for single-nutrient competition in continuous

606   cultures of micro-organisms. *SIAM Journal on Applied Mathematics*, 32(2), 366-383.

607 46. Saikaly, P. E., and Oerther, D. B. (2011). Diversity of dominant bacterial taxa in activated sludge promotes functional
608   resistance following toxic shock loading. *Microbial Ecology*, 61(3), 557-567.

609 47. Habermacher, J., Benetti, A. D., Derlon, N., and Morgenroth, E. (2015). The effect of different aeration conditions in
610   activated sludge–side-stream system on sludge production, sludge degradation rates, active biomass and extracellular
611   polymeric substances. *Water Research*, 85, 46-56.

612 48. Martínez-García, C. G., Fall, C., and Olguín, M. T. (2016). Activated sludge mass reduction and biodegradability of the
613   endogenous residues by digestion under different aerobic to anaerobic conditions: comparison and modeling.
614   *Bioresource Technology*, 203, 32-41.

615 49. Friedrich, M., and Takács, I. (2013). A new interpretation of endogenous respiration profiles for the evaluation of the
616   endogenous decay rate of heterotrophic biomass in activated sludge. *Water Research*, 47(15), 5639-5646.

617 50. Friedrich, M., Jimenez, J., Pruden, A., Miller, J. H., Metch, J., and Takács, I. (2017). Rethinking growth and decay
618   kinetics in activated sludge–towards a new adaptive kinetics approach. *Water Science and Technology*, 75(3), 501-506.

619 51. Soetaert, K. E. R., Petzoldt, T., and Setzer, R. W. (2010). Solving differential equations in R: package deSolve. *Journal
620   of Statistical Software*, 33.

621 52. Duan, L., Moreno-Andrade, I., Huang, C. L., Xia, S., and Hermanowicz, S. W. (2009). Effects of short solids retention
622   time on microbial community in a membrane bioreactor. *Bioresource Technology*, 100(14), 3489-3496.

623 53. Saikaly, P. E., Stroot, P. G., and Oerther, D. B. (2005). Use of 16S rRNA gene terminal restriction fragment analysis to
624   assess the impact of solids retention time on the bacterial diversity of activated sludge. *Applied and Environmental
625   Microbiology*, 71(10), 5814-5822.

626 54. Gonzalez-Martinez, A., Rodriguez-Sanchez, A., Lotti, T., Garcia-Ruiz, M. J., Osorio, F., Gonzalez-Lopez, J., et al.
627   (2016). Comparison of bacterial communities of conventional and A-stage activated sludge systems. *Scientific Reports*,
628   6, 18786.

629 55. Zhang, T., Shao, M. F., and Ye, L. (2011). 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14
630   sewage treatment plants. *The ISME Journal*, 6(6), 1137-1147.

631 56. Saunders, A. M., Albertsen, M., Vollertsen, J., and Nielsen, P. H. (2016). The activated sludge ecosystem contains a
632   core community of abundant organisms. *The ISME Journal*, 10(1), 11-20.

633 57. Pollice, A., Tandoi, V., and Lestingi, C. (2002). Influence of aeration and sludge retention time on ammonium oxidation
634   to nitrite and nitrate. *Water Research*, 36(10), 2541-2546.

635 58. Jiang, L. (2007). Negative selection effects suppress relationships between bacterial diversity and ecosystem
636   functioning. *Ecology*, 88(5), 1075-1085.

637 59. Prosser, J. I., Head, I. M., and Stein, L. Y. (2014). The family *Nitrosomonadaceae*. In The Prokaryotes (pp. 901-918).
638   Springer Berlin Heidelberg.

639   60. Vuono, D. C., Regnery, J., Li, D., Jones, Z. L., Holloway, R. W., and Drewes, J. E. (2016). rRNA gene expression of
640   abundant and rare activated-sludge microorganisms and growth rate induced micropollutant removal. *Environmental*
641   *Science & Technology*, 50(12), 6299-6309.

642   61. Grime, J. P., & Pierce, S. (2012). *The evolutionary strategies that shape ecosystems*. John Wiley & Sons.

643   62. Ho, A., Di Lonardo, D. P., & Bodelier, P. L. (2017). Revisiting life strategy concepts in environmental microbial
644   ecology. *FEMS Microbiology Ecology*, 93(3).

645   63. Andrews, J. H., & Harris, R. F. (1986). r-and K-selection and microbial ecology. In *Advances in Microbial Ecology* (pp.
646   99-147). Springer, Boston, MA.

647   64. Větrovský, T., and Baldrian, P. (2013). The variability of the 16S rRNA gene in bacterial genomes and its consequences
648   for bacterial community analyses. *PloS One*, 8(2), e57923.

649   65. Chandran, K., and Smets, B. F. (2001). Estimating biomass yield coefficients for autotrophic ammonia and nitrite
650   oxidation from batch respirograms. *Water Research*, 35(13), 3153-3156.

651   66. Friedrich, M., Takács, I., and Tränckner, J. (2015). Physiological adaptation of growth kinetics in activated sludge.
652   *Water Research*, 85(15), 22-30.

653   67. Vuono, D. C., Munakata-Marr, J., Spear, J. R., and Drewes, J. E. (2016). Disturbance opens recruitment sites for
654   bacterial colonization in activated sludge. *Environmental Microbiology*, 18(1), 87-99.

655   68. Blazewicz, S. J., Barnard, R. L., Daly, R. A., and Firestone, M. K. (2013). Evaluating rRNA as an indicator of microbial
656   activity in environmental communities: limitations and uses. *The ISME Journal*, 7(11), 2061-2068.

657   69. Yu, K., and Zhang, T. (2012). Metagenomic and metatranscriptomic analysis of microbial community structure and
658   gene expression of activated sludge. *PloS One*, 7(5), e38183.

659   70. Prosser, J. I. (2015). Dispersing misconceptions and identifying opportunities for the use of omics' in soil microbial
660   ecology. *Nature Reviews Microbiology*, 13(7), 439-446.

661   71. Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S., et al. (2011). The moderately efficient
662   enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*, 50(21), 4402-4410.

663   72. de Sousa Abreu, R., Penalva, L. O., Marcotte, E. M., and Vogel, C. (2009). Global signatures of protein and mRNA
664   expression levels. *Molecular BioSystems*, 5(12), 1512-1526.

665   73. Vogel, C., and Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and
666   transcriptomic analyses. *Nature Reviews Genetics*, 13(4), 227-232.

667   74. Daims, H., Lücker, S., and Wagner, M. (2016). A new perspective on microbes formerly known as nitrite-oxidizing
668   bacteria. *Trends in Microbiology*, 24(9), 699-712.

669   75. Salvado, H. (1994). Effect of mean cellular retention time on ciliated protozoan populations in urban wastewater
670   treatment plants based on a proposed model. *Water Research*, 28(6), 1315-1321.

List of Figure Legends:
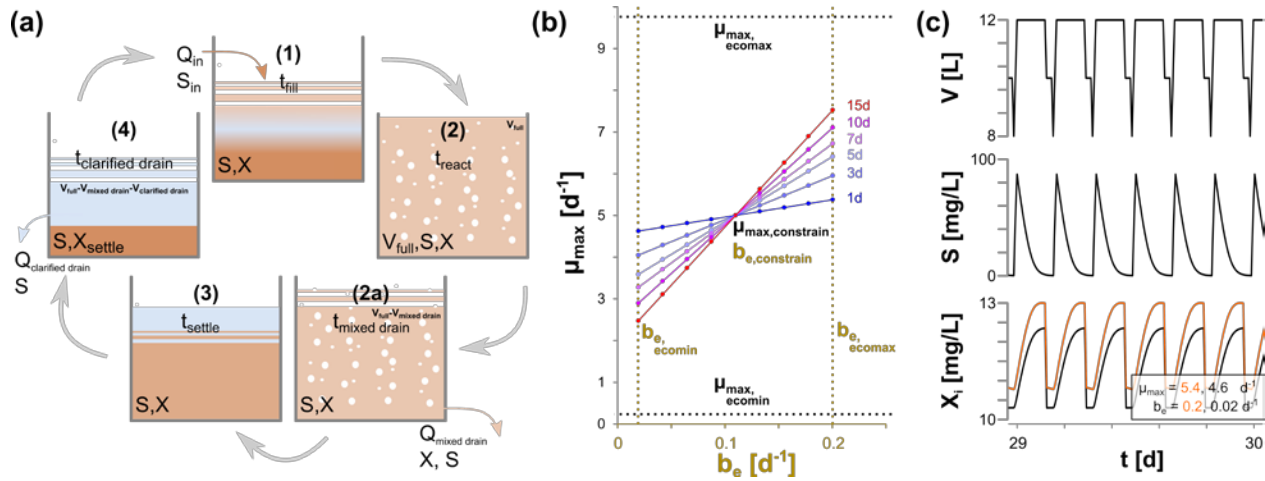
Figure 1. (a) Schematic diagram of the SBR used when running the model. All parameters are detailed in Table 1. (b) Iteratively solved 9 point persistence curves for the maximum growth rate (μ_{max}) and endogenous decay (b_e) selection range at 1, 3, 5, 7, 10, and 15 d MRT. Constraints are placed on the range of maximal growth rates ($\mu_{max\ eco}$; 0.2 to 9.8 d^{-1}) and endogenous decay constants ($b_{e\ eco}$; 0.02 to 0.2 d^{-1}), defining the ecological space available. The model uses controlling growth parameters ($\mu_{max\ constrain}$, $b_{e\ constrain}$) of 5 and 0.11 d^{-1}, respectively. (c) Example of the steady-state output for the volume, substrate concentration, and two biomasses for the 1 d MRT.
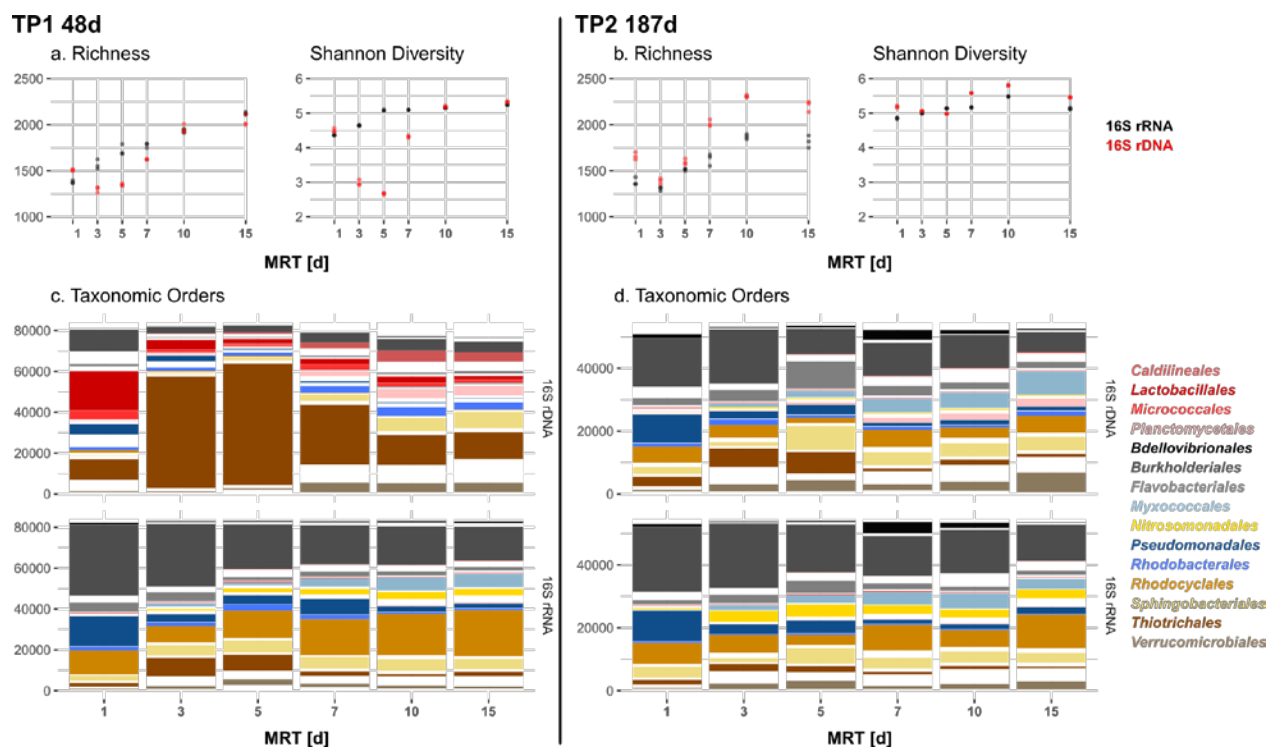
Figure 2. (a,b) The calculated diversity metrics for the rarified B2-primer amplified 16S rRNA (black) and rDNA (red) data. (c,d) The abundance data distributed into taxonomic orders; the top 10 of the sums across each time-point and source (cDNA or gDNA) were assigned a color, resulting in 15 orders being represented overall. A replicate analysis is presented for the alternate B1 primer in Supplemental Figure 15.

686

687
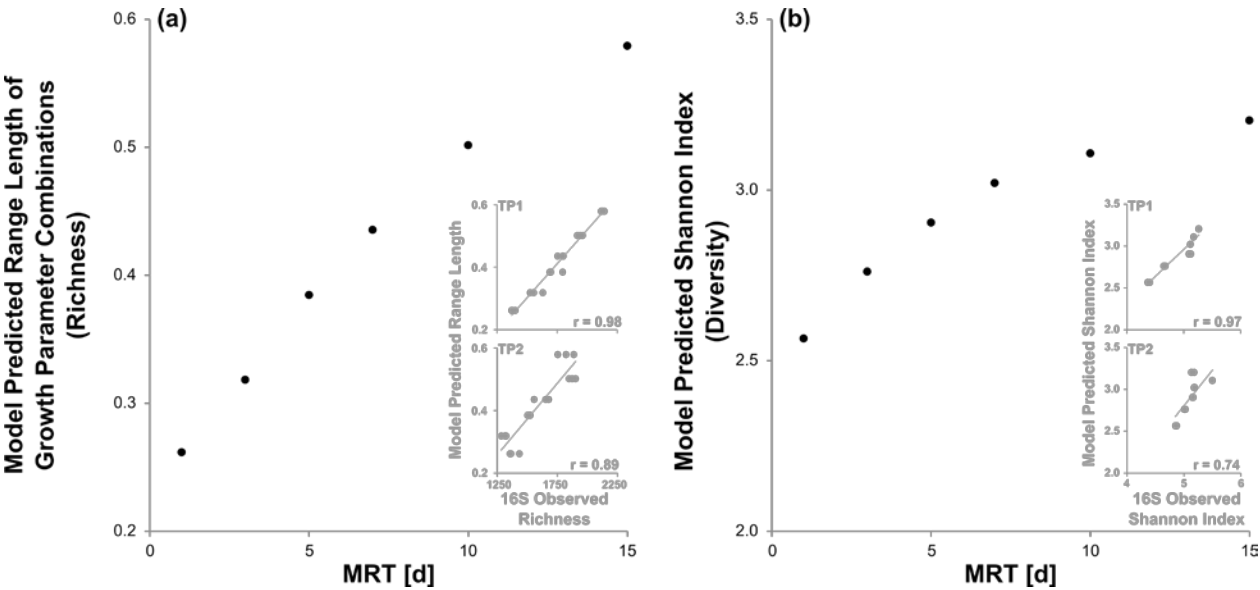
Figure 3. Alpha diversity results from the MRT-diversity model explaining the relationship between the community composition and the MRT. (a) The range of growth parameters, a proxy for the richness, with increasing MRT. The inset displays the linear relationship between the model predicted and the observed richness data. (b) The calculated Shannon diversity index. The number of species types modelled in the Shannon diversity index calculations is set to 50 times the length of the range of growth parameters. The inset displays the linear relationship between the model predicted and the observed Shannon Index data.
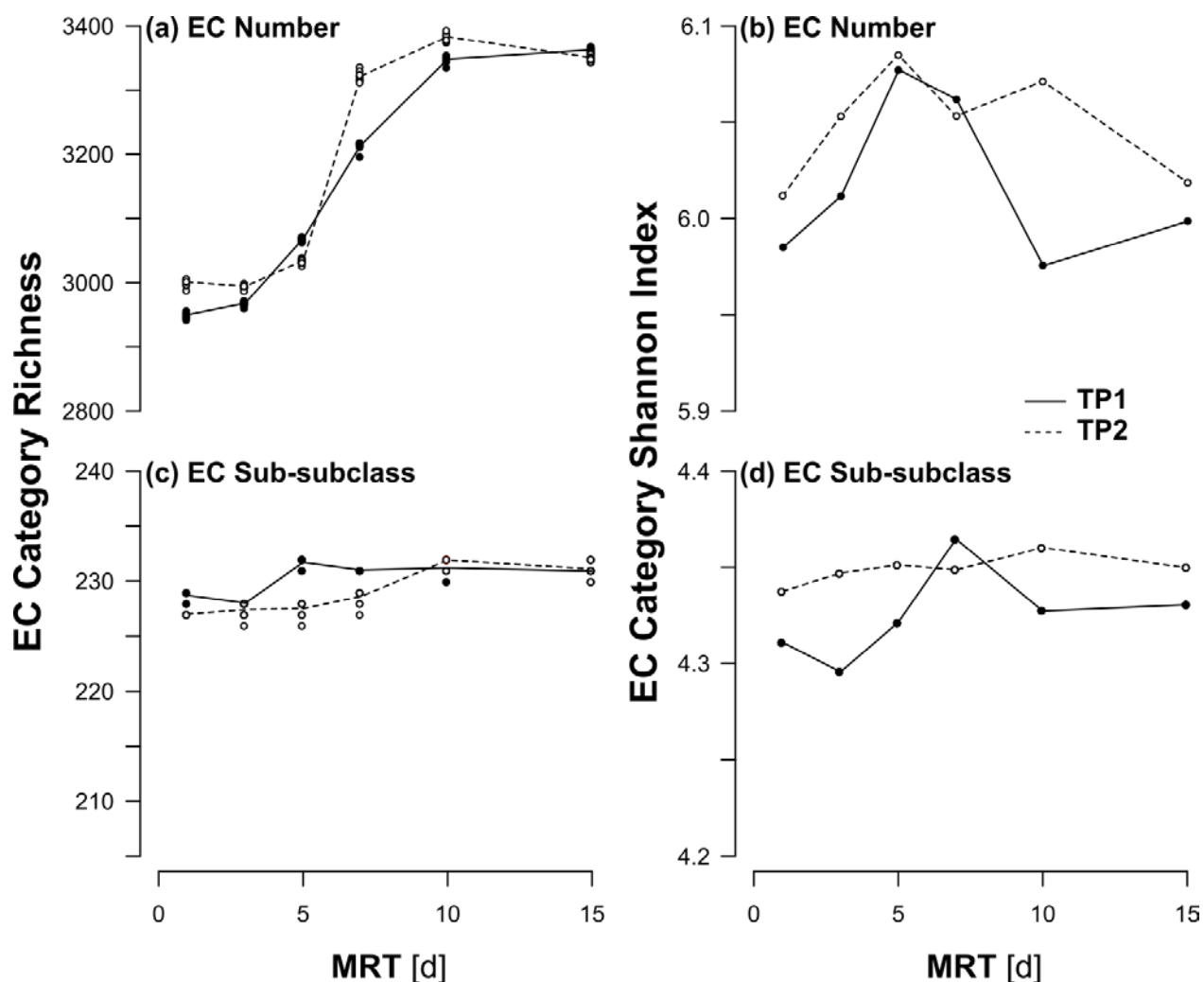
694

Figure 4. The richness (a,c) and Shannon diversity (b,d) values for the reaction-type annotated RNA data

binned into Enzyme Commission (EC) number (a,b) and sub-subclass (c,d) for TP1 (solid) and TP2

(dashed) with a fractional abundance cutoff of 10^{-7}. These diversity values were calculated for each reactor

using the rtk v0.2.5.4 package in R v3.5.1 for (a,c) 10 bootstrap sub-selections of the annotations or (b,d)

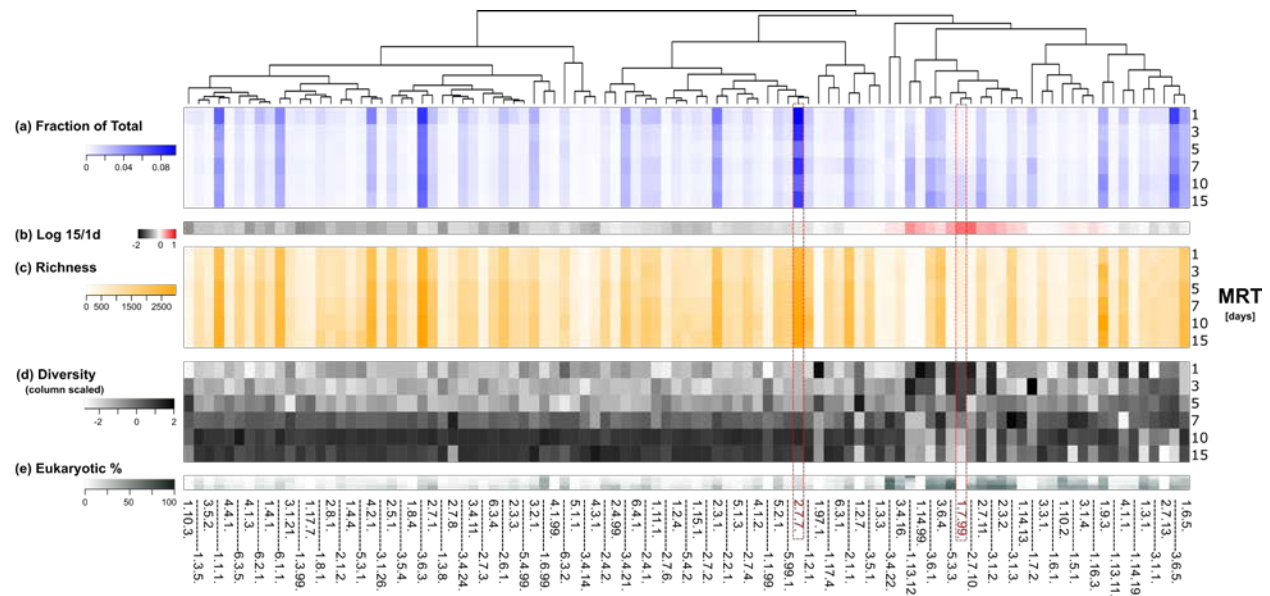the full data. The lines trace the (a,c) bootstrap mean or (b,d) calculated value.

Figure 5. Summary of the sub-subclass Enzyme Commission (EC) numbers across the MRT gradient averaged between TP1 and TP2 samples that exceed 10,000 normalized reads in at least 1 reactor (n=99). The heatmaps are organized hierarchically according to a Euclidean distance and ward clustering of the scaled EC fraction across the MRT gradient. (a) Total relative abundance. (b) The 15/1d MRT abundance log ratio. The within EC sub-subclass taxonomic (c) richness and (d) Shannon diversity metrics were calculated based on the Uniport identifiers of the organism of origin at the genus level that provided the annotation to the reads. (e) The fraction of the total reads that were annotated per EC sub-subclass originating from a Eukaryotic sequence in the Uniprot database. Note: the red dashed boxes highlight the 2.7.7 and 1.7.99 EC numbers. The full sub-subclass and top 50 EC numbers heatmaps for both TP1 and TP2 samples are presented in Supplemental Figure 17 and 18.

713 **Table 1.** Model nomenclature and parameter values

Symbol	Parameter	Units	Value	Source	Model Sensitivity
μ_{max}	Maximum growth rate	d^{-1}	-	-	Modelled Parameter
$\mu_{max,eco,min}$	Lowest maximum growth rate allowed	d^{-1}	0.2	[50,51]	Sensitive*
$\mu_{max,eco,max}$	Highest maximum growth rate allowed	d^{-1}	9.8	[52]	Sensitive*
$\mu_{max,constrain}$	Constraining growth rate used for modelling	d^{-1}	5	Set to range center	Sensitive*
b_e	Endogenous decay	d^{-1}	-	-	Modelled Parameter
$b_{e,eco,min}$	Lowest endogenous decay rate allowed	d^{-1}	0.02	[53]	Sensitive*
$b_{e,eco,max}$	Highest endogenous decay rate allowed	d^{-1}	0.2	[53]	Sensitive*
$b_{e,constrain}$	Constraining endogenous decay used for modelling	d^{-1}	0.11	Set to range center	Sensitive*
K_s	Half-saturation constant	mg/L	50	Typical value	Insensitive†
Y	Yield	mg X/mg S	0.3	Typical value	Insensitive
S_{in}	Influent substrate concentration	mg/L	250	Typical value	Insensitive
S	Substrate concentration in the reactor	mg/L	-	-	Modelled Parameter
X	Biomass concentration	mg/L	-	-	Modelled Parameter
X_{settle}	Settled biomass concentration	mg/L	-	-	Modelled Parameter
V_{full}	Full volume of reactor	L	12	Experimental Setup	-
V_t	Time variable volume of reactor	L	-	-	Modelled Parameter
$V_{mixed\ drain}$	Mixed volume removed, dependent on MRT	L	Variable‡	Experimental Setup	-
$V_{clarified\ drain}$	Clarified volume removed, 4 L - $V_{mixed\ drain}$	L	Variable‡	Experimental Setup	-
t_{cycle}	Total cycle time	hr	4	Experimental Setup	-
t_{react}	React time	hr	3	Experimental Setup	-
t_{settle}	Settle time	hr	0.5	Experimental Setup	-
$t_{mixed\ drain}$	Time to drain the mixed volume	hr	0.167	Experimental Setup	-
$t_{clarified\ drain}$	Time to drain the clarified volume	hr	0.167	Experimental Setup	-
t_{fill}	Fill time	hr	0.333	Experimental Setup	-
$Q_{mixed\ drain}$	Clarified drain flowrate	L/hr	$V_{mixed\ drain}/t_{mixed\ drain}$	Experimental Setup	-
$Q_{clarified\ drain}$	Clarified drain flowrate	L/hr	$V_{clarified\ drain}/t_{clarified\ drain}$	Experimental Setup	-
Q_{in}	Influent flowrate	L/hr	12	Experimental Setup	-

*See Supplemental Figure 15, †See Supplemental Figure 14, ‡See Supplemental Table 1

714