

**This item is the archived peer-reviewed author-version of:**

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

**Reference:**

Meerburg Francis A., Vlaeminck Siegfried, Roume Hugo, Seuntjens Dries, Pieper Dietmar H., Jauregui Ruy, Vilchez-Vargas Ramiro, Boon Nico.- 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 / International Association on Water Pollution Research - ISSN 0043-1354 - 100(2016), p. 137-145

Full text (Publisher's DOI): <http://dx.doi.org/doi:10.1016/J.WATRES.2016.04.076>

To cite this reference: <http://hdl.handle.net/10067/1399140151162165141>

1 Manuscript: Water Research

2

3 **High-rate activated sludge communities have a distinctly different structure**  
4 **compared to low-rate sludge communities, and are less sensitive towards**  
5 **environmental and operational variables.**

6

7

8 **Francis A. Meerburg<sup>a</sup>, Siegfried E. Vlaeminck<sup>a,b</sup>, Hugo Roume<sup>a</sup>, Dries Seuntjens<sup>a</sup>,**  
9 **Dietmar H. Pieper<sup>c</sup>, Ruy Jauregui<sup>c</sup>, Ramiro Vilchez-Vargas<sup>a</sup>, and Nico Boon<sup>a,\*</sup>**

10

11 <sup>a</sup>Laboratory of Microbial Ecology and Technology (LabMET), Ghent University, Coupure  
12 Links 653, 9000 Gent, Belgium

13 <sup>b</sup>Research group of Sustainable Energy, Air and Water Technology, University of  
14 Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Belgium

15 <sup>c</sup>Microbial Interactions and Processes Research Group, Helmholtz Centre for Infection  
16 Research, Inhoffenstr. 7, 38124 Braunschweig, Germany

17

18 \*Corresponding author. Phone: +32 9 264 59 76; fax: +32 9 264 62 48; E-mail:

19 [Nico.Boon@Ugent.be](mailto:Nico.Boon@Ugent.be).

20

21

22

## 23 **Abstract**

24 High-rate activated sludge processes allow for the recovery of organics and energy from  
25 wastewaters. These systems are operated at a short sludge retention time and high sludge-  
26 specific loading rates, which results in a higher sludge yield and better digestibility than  
27 conventional, low-rate activated sludge. Little is known about the microbial ecology of  
28 high-rate systems. In this work, we address the need for a fundamental understanding of  
29 how high-rate microbial communities differ from low-rate communities. We investigated  
30 the high-rate and low-rate communities in a sewage treatment plant in relation to  
31 environmental and operational variables over a period of ten months. We demonstrated  
32 that (1) high-rate and low-rate communities are distinctly different in terms of richness,  
33 evenness and composition, (2) high-rate community dynamics are more variable and less  
34 shaped by deterministic factors compared to low-rate communities, (3) sub-communities of  
35 continuously core and transitional members are more shaped by deterministic factors than  
36 the continuously rare members, both in high-rate and low-rate communities, and (4) high-  
37 rate community members showed a co-occurrence pattern similar to that of low-rate  
38 community members, but were less likely to be correlated to environmental and  
39 operational variables. These findings provide a basis for further optimization of high-rate  
40 systems, in order to facilitate resource recovery from wastewater.

## 41 **Keywords**

42 A-stage, AB-system, energy-neutral sewage treatment, co-occurrence network analysis,  
43 resource recovery

44

## 45 **1. Introduction**

46 Activated sludge treatment plays a central role in the management of domestic wastewater  
47 (sewage) and industrial wastewaters. While the conventional activated sludge process has  
48 proven its merits in terms of reliability and performance, it suffers from drawbacks such as  
49 high operational costs and limited potential for resource recovery. In recent years, high-rate  
50 activated sludge processes have gained attention because of their potential use for recovery  
51 of energy and organics from sewage, owing to their high sludge yields and good digestion  
52 properties (Meerburg et al. 2015, Jimenez et al. 2015). High-rate systems are typically  
53 operated at a short sludge retention time (SRT) of less than 2 days and a high sludge-  
54 specific loading rate (SLR) above 2 gram biochemical oxygen demand (BOD) per gram  
55 volatile suspended solids (VSS) per day (Böhnke et al. 1997). A number of municipal  
56 sewage treatment plants currently operate a high-rate stage prior to a conventional, low-  
57 rate stage. This two-stage system is known as the Adsorptions-Belebungsverfahren or AB-  
58 system (Böhnke 1977), and a number of AB-STPs are currently located in, but not limited  
59 to, the Netherlands (de Graaff and Roest 2012), Austria (Wett et al. 2007, Winkler et al.  
60 2008), Germany (de Graaff and Roest 2012, Haider et al. 2000), the US and Canada  
61 (Constantine et al. 2012) and China (Wenyi et al. 2006). Although not originally designed  
62 for the purpose of resource recovery (Böhnke et al. 1997), these STPs show great potential  
63 to improve their net energy balance, and the presence of a high-rate activated sludge is a  
64 critical factor that made the Strass STP (Austria) one of the few in the world able to  
65 achieve net energy neutrality (Wett et al. 2007). In temperate and colder climates, high-rate  
66 activated sludge treatment may be the most economically viable technology to achieve up-  
67 concentration of organics from sewage for subsequent recovery (Verstraete et al. 2009,  
68 Verstraete and Vlaeminck 2011).

69 Despite ever-improving process control, many sewage treatment plants (STPs) still  
70 struggle with operational problems that may coincide with changes in the microbial  
71 community (Gentile et al. 2007, Briones and Raskin 2003). Both for high-rate as well as  
72 conventional, low-rate systems, there is a need for better knowledge of the activated sludge  
73 community in relation to its dynamics, functional output and sensitivity toward external  
74 factors, such as changes in environmental conditions. With the development of advanced  
75 molecular techniques, a number of studies has monitored the community dynamics of  
76 activated sludge over relatively long time periods, and explored interactions of microbial  
77 species with environmental factors, with other microbial species, and with the functional  
78 output of the system (Ju and Zhang 2014, Valentin-Vargas et al. 2012, Ofițeru et al. 2010).

79 In microbial ecology, the traditional niche theory holds that microbial communities are  
80 shaped by deterministic – i.e., predictable – factors, such as environmental conditions  
81 (Chase and Leibold 2003). Changes in, for example, temperature, can have a determined  
82 influence on a species' growth rate. Different species may have different 'niches' or  
83 combinations of environmental conditions that are optimal for their growth. Thus,  
84 according to the niche theory, changes in environmental conditions will cause a shift in  
85 microbial community structure in a deterministic manner. This niche theory has been  
86 challenged by the concept of neutral change, which is based on the theory of island  
87 biogeography with a dynamic equilibrium between extinction and colonization (Hubbell  
88 2001). According to the theory of neutral community assembly, changes in microbial  
89 communities primarily reflect 'stochastic' or chance-driven processes. In other words,  
90 species may enter or disappear from a community as a result of natural fluctuations of their  
91 abundance over time, without underlying influences of environmental conditions. Recent  
92 studies suggest that activated sludge communities are shaped by both deterministic and  
93 neutral factors (Valentin-Vargas et al. 2012, Ofițeru et al. 2010, Ayarza and Erijman 2011).

94 Microbial communities are generally composed of a relatively small number of abundant  
95 species and a large number of rare species (Sogin et al. 2006). It is theorized that abundant  
96 species play a functional role in the ecosystem, while rare species merely act as a ‘seed  
97 bank’, i.e., a reserve of species present at low abundances and low activities that may  
98 become more abundant and active when conditions change (Pedros-Alio 2012). However,  
99 this may not be a general rule. For example, certain nitrifiers have been found in activated  
100 sludge at low abundance based on DNA concentrations, despite high transcription activity  
101 of nitrification-associated genes (Yu and Zhang 2012). Previous research has found that  
102 abundant sub-communities in activated sludge are less diverse than rare sub-communities  
103 and have lower species turnover rates, as indicated by the average number of new species  
104 entering the respective sub-communities per unit of time (Kim et al. 2013). However, little  
105 is known about differences in species-species and species-environment interactions  
106 between abundant and rare sub-communities.

107 While gradual progress is made in understanding the microbial ecology of conventional  
108 activated sludge systems, a large knowledge gap exists concerning high-rate activated  
109 sludge communities and their structure, dynamics, and sensitivity towards environmental  
110 factors. In this work, the high-rate and low-rate activated sludge communities of a two-  
111 stage STP were studied, and systematically compared over a period of 10 months. This  
112 work addresses four questions concerning differences in microbial ecology between high-  
113 rate and low-rate systems: (1) Are high-rate and low-rate systems distinctly different in  
114 terms of community structure? (2) Are high-rate community dynamics more variable and  
115 less governed by deterministic factors compared to low-rate communities? (3) Are  
116 community shifts in abundant and transitional sub-communities more deterministic than  
117 shifts in rare sub-communities? And (4) do high-rate community members show a lower  
118 co-occurrence and lower correlation with environmental variables than low-rate

119 community members?

120

## 121 **2. Material and Methods**

### 122 **2.1. Plant description and sampling**

123 The Nieuwveer STP in Breda (The Netherlands) operates an AB-process, and treats  
124 combined domestic and industrial wastewater from Breda and neighboring municipalities.

125 The plant was designed for a capacity of 400,000 population equivalents and the average  
126 influent flow rate during the study period was  $80,100 \text{ m}^3 \text{ d}^{-1}$ . The high-rate stage consists  
127 of a  $3,500 \text{ m}^3$  basin with an anoxic, a facultative oxic and an oxic segment. The low-rate  
128 stage treats the high-rate effluent. It consists of four parallel basins, of which the first three  
129 have a volume of  $5,400 \text{ m}^3$  and a segment train of one anoxic, two facultative oxic, two  
130 oxic and again one facultative oxic segment. The fourth basin has a volume of  $12,000 \text{ m}^3$   
131 and a segment train of two anoxic, four facultative oxic and four oxic segments. The high-  
132 rate and low-rate stages have a separate sludge recycle, each with a designed sludge  
133 recycle ratio ( $Q_{\text{recycled}} Q_{\text{influent}}^{-1}$ ) of 0.5. At the time of the study, final effluent was  
134 recirculated back to the plant inlet for improved denitrification, with a measured effluent  
135 recirculation ratio ( $Q_{\text{recirculated}} Q_{\text{influent}}^{-1}$ ) between 0.1 and 3.6. From October 2013 to July  
136 2014, near-weekly sludge samples (60 mL) were taken from the sludge recycle stream of  
137 the high-rate system and from the first segment of the largest low-rate basin. It was  
138 assumed that the sludge communities were homogenous within each system. Samples were  
139 immediately centrifuged (10 min at  $4,000g$ ). After manual homogenization of the pellets,  
140 subsamples of 0.5 mL pelletized sludge were frozen at  $-20^\circ\text{C}$  for transport and stored at -  
141  $80^\circ\text{C}$  until further processing. In parallel, fresh suspended sludge samples (1 L) were

142 transported to the lab for additional analyses within 24 h.

143

## 144 **2.2. Environmental and operational data**

145 Environmental and operational data were obtained from Waterschap Brabantse Delta (The  
146 Netherlands), who manage the STP. Total suspended solids (TSS), VSS, chemical oxygen  
147 demand (COD), BOD, sludge volume index (SVI), nitrite, nitrate, Kjeldahl nitrogen (KjN)  
148 and phosphorus concentrations were determined by Waterschap Brabantse Delta according  
149 to standard methods (Greenberg et al. 1992). Volume-weighted average diameters ( $D_{4,3}$ ) of  
150 the sludge flocs were measured with a Mastersizer S (Malvern, Malvern, UK), as described  
151 by Courtens et al. (2014). Extracellular polymeric substances (EPS) were extracted from  
152 the sludge flocs using a heat extraction protocol described by Judd and Judd (2006) and  
153 subsequently stored at -20°C. For determination of the EPS protein content, samples were  
154 alkalified to a final concentration of 1 M NaOH, and analyzed using the Lowry protein  
155 assay (Lowry et al. 1951) with bovine serum albumin as a standard.

156 Data collection of environmental and operational variables did not always coincide with  
157 sampling of the microbial communities. For continuously measured variables such as  
158 temperature, recirculation factor, hydraulic residence time, oxygen concentrations and  
159 rainfall, average values were taken for a two-day interval before each sludge sample. For  
160 the intermittently measured variables, the value closest in time to each sludge sample was  
161 used within a range of a few days before to 1 day after sludge sampling. **Table 1** lists all  
162 environmental and operational variables used in this study, and their abbreviations.

## 163 **2.3. Community analysis**

164 DNA was extracted from the pelletized sludge samples using a FastPrep-24 system (MP  
165 Biomedicals, California, USA), and precipitated according to the protocol described by



166 Vilchez-Vargas et al. (2013). The DNA pellets were resuspended in 100  $\mu$ L MilliQ water.  
167 The quantity of the DNA was tested by monitoring the absorbance at 260 nm and  
168 absorbance ratios at 260 nm and 280 nm using a NanoDrop ND-1000 (Thermo Scientific,  
169 Massachusetts, USA), and the quality was checked by electrophoresis on a 1% (w/v)  
170 agarose gel. Samples were sequenced using the high-throughput MiSeq Illumina platform  
171 (Illumina, California, USA). Regions V5-V6 of the 16S rRNA gene were amplified, and  
172 targeted with adapters and barcodes suitable for Illumina sequencing, as previously  
173 described (Bohorquez et al. 2012, Camarinha-Silva et al. 2014). Quality filtering was  
174 performed as described by Camarinha-Silva et al. (2014). Read length was between 140  
175 and 273 nucleotides. Reads were clustered using the Mothur pipeline (Schloss et al. 2009),  
176 allowing two mismatches. This resulted in 1,677 unique taxa (phylotypes). The phyloseq  
177 package (McMurdie and Holmes 2013) was used in R (version 3.0.2) to randomly  
178 normalize each sample to the minimum sequencing depth of 15,186 reads, and the vegan  
179 package (Oksanen et al. 2013) was used to visualize that all samples reached a plateau in  
180 the rarefaction curve (**Supplementary Figure S1**). Phylotypes were annotated in the RDP  
181 classifier (Cole et al. 2014) using the naïve Bayesian classification (Wang et al. 2007) with  
182 a threshold of 80%, and manually analyzed using the seqmatch function. A taxonomic  
183 level was only assigned when 16S rRNA gene fragments of previously described isolates  
184 or uncultured representatives of that taxon showed  $\leq 2$  mismatches. Sequences were  
185 deposited in the European Nucleotide Archive (accession numbers LT217663 to  
186 LT219428).

187

#### 188 **2.4. Statistical analysis**

189 Statistical comparisons of community indices (richness, evenness, dynamics and relative

190 phylum abundance) between the high-rate and low-rate systems were performed in R. The  
191 Shapiro-Wilk test was used to test the normality of the data residuals. The null hypothesis  
192 of normality was rejected for the evenness and dynamics of the high-rate system, and for  
193 some of the relative phylum abundances in the high-rate and low-rate systems. Therefore,  
194 pairwise statistical comparisons of community indices between the high-rate and low-rate  
195 systems were performed using the Mann-Whitney  $U$  test as a non-parametric alternative  
196 for the Student's  $t$ -test. Differences were considered significant at a  $p$ -value below 0.05.  
197 Ordination and calculation of diversity and dissimilarity indices were performed using the  
198 vegan package in R. Unimodal ordination methods (correspondence analysis, CA; and  
199 canonical correspondence analysis, CCA) were preferred, since the gradient lengths of the  
200 detrended correspondence analyses were always  $\approx 4$  (Ramette 2007). For all ordinations,  
201 only environmental variables that significantly correlated to the unconstrained CA axes  
202 (9999 permutations) were considered for variation partitioning in CCA analysis. Pearson  
203 and Spearman correlations were calculated using the hmisc package in R (Harrell 2014).  
204 To construct co-occurrence networks, the absolute phylotype (Phy) abundance matrices  
205 were used to calculate Pearson correlations in a pair-wise manner. Only significant  
206 correlations above 0.65 were used for network construction. The undirected network was  
207 visualized and analyzed using Cytoscape (version 3.2.1) (Shannon et al. 2003), using an  
208 organic layout.

209

## 210 **3. Results and Discussion**

### 211 **3.1. Question 1**

212 **“Are high-rate and low-rate systems distinctly different in terms of community**

213 **structure?”**

214

215 A total of 22 environmental variables were monitored for the high-rate and 19 for the low-  
216 rate systems of the sewage treatment plant (**Table 1**). The main differences between the  
217 two systems were the incoming BOD concentration, the SLR and the  $D_{4,3}$ , which were  
218 considerably higher in the high-rate system, and the HRT and SRT, which were  
219 considerably shorter. Throughout the study period, no major disruptions of plant  
220 performance occurred, and the STP was able to remove 85-96% of COD and 95-99% of  
221 TSS. Removal performances of nitrogen (42-91%) and phosphorus (33-95%) were more  
222 variable, with minima occurring between the colder months of November 2013 to February  
223 2014.

224 CA ordination of the phylotype-sample abundance matrices showed a clear separation  
225 between samples of the high-rate and low-rate systems along the primary ordination axis,  
226 while the secondary axis showed variation within each stage. A major fraction of the  
227 phylotypes also clustered according to a similar pattern (**Figure 1**). Fitted environmental  
228 variables indicate the direction of each variable across the ordination space, and their  
229 length reflects the strength of correlation to the ordination axes. The distinction between  
230 samples and phylotypes along the first ordination axis was most strongly correlated to the  
231 environmental variables of  $HRT_{nom}$ ,  $SRT_{syst}$ , SLR, COD/N ratio,  $D_{4,3}$ , SVI, BOD and TSS.  
232 Variation along the second ordination axis was most strongly correlated to the time.

233 Over the entire sampling period, 266 phylotypes were detected only in the high-rate, 990  
234 only in the low-rate and 510 phylotypes were detected at least once in both stages.

235 Community-wide comparison showed that the high-rate system had a considerably lower  
236 observed richness ( $289 \pm 48$  phylotypes) and Pielou's evenness ( $0.62 \pm 0.06$ ), compared to  
237 the low-rate system ( $668 \pm 63$  phylotypes and  $0.82 \pm 0.02$ , respectively) (**Supplementary**  
10

238 **Figure S2**), and these differences were highly significant ( $p < 10^{-12}$ ). These results are  
239 complementary to a recent study of ten single-time-point samples from different high-rate  
240 and low-rate STPs (Gonzalez-Martinez et al. 2016), which showed that, of the five studied  
241 environmental variables, the SRT and HRT were most strongly correlated with differences  
242 in microbial community structure. However, mentioned study did not incorporate several  
243 environmental factors that were shown in current study to associate with differences in  
244 microbial community structure between high-rate and low-rate activated sludge (see  
245 above), including time. Gonzalez-Martinez et al. (2016) also demonstrated that the  
246 microbial communities of the high-rate sludge plants were consistently less diverse than  
247 the low-rate communities. Saikaly and Oerther (2004), argued that species richness  
248 increases with SRT. However, experimental studies on membrane bioreactors (MBRs)  
249 have demonstrated positive (Duan et al. 2009), negative (Saikaly et al. 2005) and neutral  
250 effects (Bagchi et al. 2015, Tan et al. 2008, Teksoy Başaran et al. 2014) of SRTs between  
251 0.5 and 33 d on community richness and evenness. Besides the SRT, the evenness in the  
252 low-rate reactor may also explain its higher species richness, since systems with higher  
253 evenness are theorized to provide more niche space for microbial colonization (van der  
254 Gast et al. 2006). A study on two full-scale sewage treatment plants with large differences  
255 in SRT and SLR showed that samples from the two reactors clustered separately in CCA,  
256 and that differences in community composition could be correlated to the SRT, SLR, HRT  
257 and temperature (Valentin-Vargas et al. 2012). Neutral factors are also known to influence  
258 activated sludge communities (Valentin-Vargas et al. 2012, van der Gast et al. 2008).  
259 Nonetheless, random factors alone cannot explain the differences in the sludge  
260 communities described in this study, considering that the hydraulic connection of the two  
261 systems creates a continuous cross-inoculation, and that differences in community  
262 structure are pronounced and consistent over time. This raises the question as to how

263 community structure and function of high-rate and low-rate systems are affected when a  
264 substantial amount of biomass is continuously transferred from one system to another, as is  
265 the case in the Hybrid® process (Winkler et al. 2004). To exploit the full capacity of a two-  
266 stage STP, one may argue that it is essential that both stages have distinctly different  
267 microbial communities to be better adapted to the specific purpose of each stage. In this  
268 study, it was clear that the community structure and composition of the high-rate and low-  
269 rate systems were distinctly and consistently different, and that this could be attributed to  
270 differences in operational and environmental factors.

271

### 272 *3.2. Question 2*

273 **“Are high-rate community dynamics more variable and less governed by**  
274 **deterministic factors compared to low-rate communities?”**

275

276 The observed community dynamics were expressed as dissimilarity between consecutive  
277 samples in a moving-window approach with a fixed one-week interval (**Figure 2**). The  
278 high-rate system experienced an alternation between periods of stronger changes and more  
279 stable periods, whereas the low-rate system displayed a more consistent level of dynamics  
280 over time. Remarkably, the average dynamics in the two systems was similar, with a  
281 weekly Bray-Curtis dissimilarity of  $0.19 \pm 0.06$  in the high-rate system and  $0.20 \pm 0.03$  in  
282 the low-rate system ( $p > 0.05$ ). At short SRT, and thus high specific growth rate, it has  
283 been suggested that sludge systems experience a higher degree of dynamics, due to  
284 oscillations in population abundances (Saikaly and Oerther 2004, Curtis et al. 2003) and a  
285 number of studies has found a correlation between short SRT and higher community  
286 dynamics (Valentin-Vargas et al. 2012, Duan et al. 2009). On the other hand, systems with

287 a higher diversity are thought to harbor more redundancy within functional groups  
288 (Briones and Raskin 2003), and richer systems may therefore experience dynamic  
289 population changes without affecting functional stability. Possibly, the similar degree of  
290 dynamics for the high-rate and low-rate systems in this study was a result of the conflicting  
291 effects of SRT and diversity on system dynamics.

292 The taxa-time relationship describes the accumulation of new phylotypes over time, and  
293 may be explained by fitting a power-law function:

$$294 \quad S = c T^w \quad \text{(Equation 1)}$$

295 where  $S$  is the cumulative number of taxa over time  $T$ ,  $c$  is a constant and  $w$  is the temporal  
296 scaling exponent (Preston 1960), which is a measure of relative species turnover rate. The  
297 temporal scaling exponents for the high-rate (0.262,  $R^2 = 0.960$ ) and low-rate system  
298 (0.249,  $R^2 = 0.968$ ) were similar ( $p > 0.05$ ) (**Supplementary Figure S3**), and fell within  
299 the lower range of values between 0.21 – 0.50 reported for activated sludge systems (Kim  
300 et al. 2013, Wells et al. 2011, Shade et al. 2013, Hai et al. 2014, Ibarbalz et al. 2014). The  
301 similarity of temporal scaling exponents of the high- and low-rate community is  
302 unexpected, given that these systems differed in species richness and selective pressure  
303 caused by differences in SRT. For example, Ayarza and Erijman (2011) found that  
304 activated sludge communities with a more diverse initial richness experienced higher  
305 species turnover rates. In contrast, van der Gast et al. (2008) reported lower turnover rates  
306 as activated sludge communities experienced a higher selective pressure. In this work, the  
307 high-rate system had a lower species richness, which would be expected to lead to lower  
308 turnover rates. Additionally, the high-rate system had a higher selective pressure on  
309 microbial growth rates because of the shorter SRT, which would also be expected to lead  
310 to lower turnover rates. The fact that community dynamics and relative species turnover  
311 rate were very similar in the high-rate and low-rate systems may therefore indicate that

312 other factors exist, besides species richness and SRT, that influence community turnover  
313 rates, and that were not included in this study.

314 To quantify the relative importance of deterministic factors shaping the overall community  
315 structure, variation partitioning was performed by CCA ordination of the high-rate and  
316 low-rate communities separately (**Table 2**). Note that time may not be a true environmental  
317 factor, and community changes over time may reflect deterministic as well as neutral  
318 changes (Lynch and Neufeld 2015).

319 Assuming that this study included the environmental variables most relevant for the  
320 ecology of activated sludge communities (Valentin-Vargas et al. 2012, Wells et al. 2011,  
321 Hai et al. 2014, Ibarbalz et al. 2014), the percentage of unexplained variation was 52.5% in  
322 the high-rate and 44.1% in the low-rate system. This suggests that high-rate activated  
323 sludge communities are more shaped by neutral factors than low-rate communities. As a  
324 consequence, high-rate systems may potentially be less controllable for technological  
325 applications, but also less subject to disturbance from environmental perturbations.

326

### 327 **3.3. Question 3**

328 **“Are community shifts in abundant and transitional sub-communities more**  
329 **deterministic than shifts in rare sub-communities?”**

330

331 The threshold of abundance to distinguish between abundant and rare members has been  
332 arbitrarily set at values from 0.01 % to 1 % of the total community (Pedros-Alio 2012, Kim  
333 et al. 2013, Bagchi et al. 2015, Campbell et al. 2011). For any given dataset, it is important  
334 to assess the impact of varying this threshold, because it may influence the results of  
335 further ecological analyses (Gobet et al. 2010). In this work, the threshold of distinction

336 between abundant and rare community members was varied between 0.01% and 1% and  
337 the distribution between continuously abundant, transitional and continuously rare  
338 phylotypes in both datasets was evaluated (**Supplementary Figure S4**). A threshold of  
339 0.1% relative abundance was considered to yield the most informative distribution: in the  
340 high-rate system, this threshold resulted in a continuously abundant sub-community of  
341 1.7% of phylotypes and 60.7% of all sequences, and a continuously rare sub-community of  
342 67% of phylotypes and 3.3% of sequences, with the remainder constituting the transitional  
343 sub-community. In the low-rate system, a similar distribution was obtained (**Table 3**).  
344 The distribution of phyla differed along sub-communities. In all cases, Proteobacteria were  
345 dominant, followed by Bacteroidetes. In both the high-rate and the low-rate system, the  
346 continuously abundant sub-communities were nearly completely composed of  
347 Proteobacteria while the transitional sub-communities were near-equally dominated by  
348 Proteobacteria and Bacteroidetes. The continuously rare sub-communities were again  
349 dominated by Proteobacteria, followed by Bacteroidetes and a number of other phyla  
350 (**Supplementary Figure S5**). A similar dominance of Proteobacteria and, to a lesser  
351 extent, Bacteroidetes was also reported in other studies that described phylogenetic  
352 distributions in abundant, transitional and/or rare sub-communities of activated sludge (Ju  
353 and Zhang 2014, Kim et al. 2013, Ibarbalz et al. 2014, Ju et al. 2014, Shade et al. 2014,  
354 Saunders et al. 2016), and the dominance of Proteobacteria and Bacteroidetes has been  
355 observed in both high-rate and low-rate activated sludge communities (Gonzalez-Martinez  
356 et al. 2016). Still, significant differences were found for the relative abundance of  
357 Proteobacteria and Bacteroidetes between each of the sub-communities of the high-rate  
358 and low-rate system ( $p$ -value  $< 10^{-3}$  for each pairwise comparison). This suggests that these  
359 phyla play different functional roles in the system. For example, the lower relative  
360 abundance of Bacteroidetes in the abundant sub-communities compared to the transitional



361 sub-communities raises the question whether species of this phylum are less likely to exert  
362 a core ecosystem function.

363 From the assumed functional roles of the abundant and transitional sub-communities, it  
364 may be hypothesized that dynamic changes in these sub-communities are more  
365 deterministic than changes in the rare sub-community. A similar phenomenon has also  
366 been observed in macroecological studies, where the relative abundance of core species  
367 relied more on biological factors, while satellite species were more determined by random  
368 dispersal (Magurran and Henderson 2003, Ulrich and Zalewski 2006). Separate CA  
369 analyses for each sub-community of the high-rate and low-rate system were performed  
370 (**Supplementary Figure S6**). Subsequent CCA analyses showed that, in both the high-rate  
371 and low-rate systems, larger fractions of community variation could be correlated to  
372 changes of environmental variables for the abundant and transitional sub-communities than  
373 for the continuously rare sub-communities (**Table 2**). The same trend was observed when  
374 different abundance thresholds were used to distinguish the sub-communities from one  
375 another. Indeed, as reviewed by Lynch and Neufeld (2015), previous studies on aquatic  
376 ecosystems have shown that rare sub-communities may be disproportionately influenced by  
377 random factors, but may retain a certain degree of activity and susceptibility to selective  
378 environmental factors. The results of this study support the theory that part of the rare  
379 community may act as a ‘seed bank’ waiting for the right growth conditions, and  
380 controlled by neutral factors.

381

#### 382 **3.4. Question 4**

383 **“Do high-rate community members show a lower co-occurrence and lower**  
384 **correlation with environmental variables than low-rate community members?”**

385

386 Microbial co-occurrence may be direct (e.g., biological interactions) or indirect (e.g.,  
387 shared ecological niches), but always reflect a deterministic relationship, rather than  
388 neutral association (Barberan et al. 2012). Co-occurrence networks of the high-rate and  
389 low-rate communities were created, based on pairwise Pearson correlations between  
390 phylotype abundances (**Figure 3**). The continuously rare sub-communities were excluded  
391 from the network analysis to filter out infrequent phylotypes, and to avoid that the network  
392 loses specificity due to low site similarities (Berry and Widder 2014). After their exclusion  
393 from the datasets, the mean Jaccard similarity between sites was 49% for the high-rate  
394 system and 47% for the low-rate system, and thus higher than the minimum of 20%  
395 recommended by Berry and Widder (2014).

396 The average node degree – i.e., the average number of connections per node – was 9.4 in  
397 the high-rate network and 18.5 in the low-rate network. This means that both systems may  
398 be considered highly interconnected (Barberan et al. 2012). With 256 nodes, the high-rate  
399 network had 1203 edges, which constituted 3.7% of the total of  $3.3 \times 10^4$  possible edges of a  
400 fully saturated network. The low-rate network had 581 phylotypes and 5,378 edges, which  
401 constituted 3.2% of the total of  $1.7 \times 10^5$  possible edges. Therefore, when corrected for the  
402 number of network nodes, the high-rate and low-rate communities had a similar co-  
403 occurrence pattern. In the high-rate network, five loosely connected clusters of nodes could  
404 be distinguished, and in the low-rate network three. Throughout the study period, these  
405 clusters successively dominated their respective community in terms of abundance  
406 (**Supplementary Figure S7**).

407 Keystone community members are defined as having a disproportionately strong effect on  
408 their ecosystem functioning relative to their abundance (Paine 1995). To identify keystone

409 members from a co-occurrence network, the most likely candidates are nodes that are  
410 highly connected and centrally clustered, and can be indicated by network metrics, such as  
411 a high node degree, low betweenness centrality and high closeness centrality (Berry and  
412 Widder 2014). Based on evaluation of these three parameters, the strongest keystone  
413 characteristics were found for Comamonadaceae gen. sp. (Phy 229), Bacteroidetes gen. sp.  
414 (Phy 208), SR1 gen. sp. (Phy 313) and *Rhodoferrax* sp. (Phy 31) in the high-rate system  
415 (**Supplementary Table S1**). *Rhodoferrax* is known for its facultative photoheterotrophic  
416 and denitrifying metabolism (McIlroy et al. 2015), and showed a strong negative  
417 correlation with the HRT ( $r = -0.74$ ) and KjN concentration ( $r = -0.72$ ) in the high-rate  
418 system. In the low-rate system, the strongest keystone characteristics were found for  
419 *Sorangium* spp. (Phy 513, Phy 542 and Phy 245) (**Supplementary Table S2**). These three  
420 phylotypes showed a negative correlation with the KjN concentration ( $r = -0.73$  to  $-0.67$ ).  
421 *Sorangium* is a genus of Myxobacteria with cellulose-degrading capabilities (Hou et al.  
422 2006). No *Sorangium* sp. were detected in the high-rate system, which may be a result of  
423 their slow growth rate (Rachid et al. 2007). In both systems, all of the phylotypes with the  
424 strongest keystone characteristics belonged to the transitional sub-community, except for  
425 *Dokdonella* sp. (Phy 7), a keystone candidate in both systems, which was transitional in the  
426 high-rate system and continuously abundant in the low-rate system. *Dokdonella* is an  
427 aerobic heterotroph known for its presence in activated sludge (McIlroy et al. 2015). In the  
428 low-rate system, its abundance strongly correlated with temperature ( $r = 0.73$ ). Certain  
429 phylotypes were continuously abundant but correlated neither with any other phylotype nor  
430 with any environmental variable included in this study. In the high-rate system, these  
431 included *Acidovorax* sp. (Phy 2), a genus of aerobic and denitrifying heterotrophic bacteria  
432 (McIlroy et al. 2015), and *Aquabacterium* sp. (Phy 12), a genus of microaerophilic  
433 denitrifying bacteria that may play a role in phosphorus removal (Kalmbach et al. 1999). In

434 the low-rate system, these included Phy 2, *Sulfuritalea* sp. (Phy 14), a facultatively  
435 autotrophic genus involved in sulfur and hydrogen oxidation (Kojima and Fukui 2011),  
436 Sphingobacteriales gen. sp. (Phy 19), Chitinophagaceae gen. sp. (Phy 74), and *Derxia* sp.  
437 (Phy 101), a genus of facultatively autotrophic hydrogen oxidizers (Dworkin et al. 2006). It  
438 may be argued that the continuously abundant presence of these phylotypes through time  
439 suggests that their abundance is influenced by unidentified deterministic functional or  
440 environmental factors, rather than neutral assembly. On the other hand, previous research  
441 has demonstrated that some microorganisms may be abundant in activated sludge despite a  
442 low net growth-rate, due to the continuous influx of microorganisms with the sewage  
443 (Saunders et al. 2016).

444 To assess whether correlations with environmental variables are less strong in high-rate  
445 communities than in low-rate activated sludge communities, correlations between  
446 individual phylotypes and environmental variables were calculated, and the percentage of  
447 correlations exceeding a given threshold counted (**Figure 4**).

448 Between  $r = 0.45$  and  $r = 0.7$ , the percentage of absolute correlation coefficients exceeding  
449 a given threshold in the high-rate community was always roughly half the fraction in the  
450 low-rate community. This indicates that community members in high-rate activated sludge  
451 are less likely to be correlated to environmental variables than in low-rate activated sludge.  
452 In the high-rate community, the strongest correlations were found with time (43 phylotypes  
453 with absolute correlation coefficient  $> 0.7$ ), temperature (25 phylotypes) and KjN (5  
454 phylotypes). In the low-rate community, these were time (135 phylotypes), temperature (46  
455 phylotypes), nitrogen removal efficiency (22 phylotypes) and hydraulic retention time (11  
456 phylotypes).

457 Overall, these results confirm that high-rate community members are less strongly  
458 correlated to environmental variables than members of low-rate activated sludge

459 communities. This supports the hypothesis that high-rate communities are more subjected  
460 to neutral factors than low-rate communities, such as stronger oscillations in species  
461 abundances caused by the shorter SRT (Saikaly and Oerther 2004), as presented in  
462 **Question 2**, or continuous random colonization by new species from the influent  
463 microbiome (Ofițeru et al. 2010).

464

## 465 **4. Conclusions**

466 We investigated the microbial ecology of high-rate and low-rate activated sludge  
467 communities of a full-scale STP system, in terms of community structure, composition and  
468 sensitivity toward changes in environmental and operational variables. We showed that  
469 that:

- 470 • High-rate and low-rate communities are distinctly different in terms of richness,  
471 evenness and composition
- 472 • Both communities show a similar degree of weekly dynamics, but high-rate  
473 system dynamics are more variable
- 474 • High-rate communities are less shaped by deterministic factors, such as  
475 environmental and operational variables, than low-rate communities
- 476 • In both systems, continuously core and transitional sub-communities are more  
477 shaped by deterministic factors than the sub-community of continuously rare  
478 members
- 479 • High-rate community members show a co-occurrence pattern similar to that of  
480 low-rate community members, but are less likely to be correlated to environmental  
481 variables.

482

483 These findings provide a first basis for understanding how high-rate communities differ  
484 from conventional low-rate communities, and may facilitate a faster adoption of high-rate  
485 processes for improving the energy balance of sewage treatment plants. Differences in  
486 operational and environmental variables in a high-rate system result in a distinctly different  
487 microbial community compared to low-rate systems. This community differentiation may  
488 contribute to the improved overall performance of two-stage STPs in terms of energy and  
489 resource recovery. Additionally, the relatively high importance of neutral factors in  
490 shaping the community of high-rate systems suggest that they may be less sensitive  
491 towards external shocks and perturbations, but at the same time be more challenging to  
492 steer by controlling the operational conditions. Future studies should assess the  
493 implications for process engineering of high-rate systems, in order to develop specialized  
494 optimization and control strategies.

495

## 496 **5. Acknowledgements**

497 F.A.M and S.E.V were supported, respectively, as doctoral candidate (Aspirant) and  
498 postdoctoral fellow by the Research Foundation Flanders (FWO-Vlaanderen). H.R. was  
499 supported by the European Research Council (ERC) starter grant ELECTROTALK. D.S.  
500 was supported by the Agency for Innovation by Science and Technology in Flanders  
501 (IWT). R.V.V. and N.B. were supported by the Inter-University Attraction Pole (IUAP) ‘ $\mu$ -  
502 manager’ funded by the Belgian Science Policy (BELSPO, P7/25). We thank Arne  
503 Verliefde and Quenten Denon for assistance with the particle size measurements, Hans  
504 Mollen, Leonie Hartog, Etteke Wypkema and Waterschap Brabantse Delta for supplying  
505 plant data and allowing the use of the plant facilities, Harry Vrins and Antoon Buijnsters

506 for assistance with sampling, Bert Bundervoet and Colsen NV for practical assistance and  
507 logistics, Iris Plumeier and Silke Kahl for assistance with the sequencing, Frederiek-  
508 Maarten Kerckhof and Stijn Decubber for assistance with the analyses, and Jo De Vrieze  
509 and Ruben Props for the inspiring scientific discussions.  
510

## 511 **6. References**

- 512 Meerburg, F.A., Boon, N., Van Winckel, T., Vercamer, J.A.R., Nopens, I. and Vlaeminck,  
513 S.E. (2015) Toward energy-neutral wastewater treatment: A high-rate contact stabilization  
514 process to maximally recover sewage organics. *Bioresource Technology* 179, 373-381.  
515 Jimenez, J., Miller, M., Bott, C., Murthy, S., De Clippeleir, H. and Wett, B. (2015) High-  
516 rate activated sludge system for carbon management – Evaluation of crucial process  
517 mechanisms and design parameters. *Water Research* 87, 476-482.  
518 Böhnke, B., Diering, B. and Zuckut, S.W. (1997) Cost-effective wastewater treatment  
519 process for removal of organics and nutrients I. *Water Engineering & Management* 144(5),  
520 30-34.  
521 Böhnke, B. (1977) Das Adsorptions-Belebungsverfahren. *Korrespondenz Abwasser* 24(2),  
522 33-42.  
523 de Graaff, M. and Roest, K. (2012) Inventarisatie van AB systemen - Optimal process  
524 conditions in the A-stage. *KWR report 2012-094* (in Dutch), p. 65.  
525 Wett, B., Buchauer, K. and Fimml, C. (2007) Energy self-sufficiency as a feasible concept  
526 for wastewater treatment systems, pp. 21-24, Singapore.  
527 Winkler, S., Gasser, M., Schattle, W., Kremmel, D., Kletzmayer, P. and Matsche, N. (2008)  
528 Upgrading of wastewater treatment plants for nutrient removal under optimal use of  
529 existing structures. *Water Science and Technology* 57(9), 1437-1443.  
530 Haider, S., Vanrolleghem, P.A. and Kroiß, H. (2000) Low sludge age and its consequences  
531 for metabolisation, storage and adsorption of readily biodegradable substrate (S<sub>s</sub>), 3 - 7  
532 July 2000, Paris, France.  
533 Constantine, T., Houweling, D. and Kraemer, J.T. (2012) "Doing the two-step" - Reduced  
534 energy consumption sparks renewed interest in multistage biological treatment, 29  
535 September - 3 October 2012, New Orleans, LA, U.S.A.  
536 Wenyi, D., Hong, D., Li-an, Z., Jia, M. and Baozhen, W. (2006) Operational retrofits of  
537 AB process for biological removal of nitrogen and phosphorus. *Water Practice &*  
538 *Technology* 1(4).  
539 Verstraete, W., van de Caveye, P. and Diamantis, V. (2009) Maximum use of resources  
540 present in domestic "used water". *Bioresource Technology* 100(23), 5537-5545.  
541 Verstraete, W. and Vlaeminck, S.E. (2011) ZeroWasteWater: short-cycling of wastewater  
542 resources for sustainable cities of the future. *International Journal of Sustainable*  
543 *Development and World Ecology* 18(3), 253-264.  
544 Gentile, M.E., Jessup, C.M., Nyman, J.L. and Criddle, C.S. (2007) Correlation of  
545 Functional Instability and Community Dynamics in Denitrifying Dispersed-Growth  
546 Reactors. *Applied and environmental microbiology* 73(3), 680-690.

547 Briones, A. and Raskin, L. (2003) Diversity and dynamics of microbial communities in  
548 engineered environments and their implications for process stability. *Current Opinion in*  
549 *Biotechnology* 14(3), 270-276.

550 Ju, F. and Zhang, T. (2014) Bacterial assembly and temporal dynamics in activated sludge  
551 of a full-scale municipal wastewater treatment plant. *The ISME Journal*, 1-13.

552 Valentin-Vargas, A., Toro-Labrador, G. and Massol-Deya, A.A. (2012) Bacterial  
553 community dynamics in full-scale activated sludge bioreactors: Operational and ecological  
554 factors driving community assembly and performance. *PLOS one* 7(8), e42524.

555 Ofiteiru, I.D., Lunn, M., Curtis, T.P., Wells, G.F., Criddle, C.S., Francis, C.A. and Sloan,  
556 W.T. (2010) Combined niche and neutral effects in a microbial wastewater treatment  
557 community. *Proceedings of the National Academy of Sciences of the United States of*  
558 *America* 107(35), 15345-15350.

559 Chase, J.M. and Leibold, M.A. (2003) *Ecological niches: Linking classical and*  
560 *contemporary approaches*, University of Chicago Press, Chicago, IL, U.S.A.

561 Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*,  
562 Princeton University Press, Princeton, NJ, U.S.A.

563 Ayarza, J.M. and Erijman, L. (2011) Balance of neutral and deterministic components in  
564 the dynamics of activated sludge floc assembly. *Microbial Ecology* 61(3), 486-495.

565 Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta,  
566 J.M. and Herndl, G.J. (2006) Microbial diversity in the deep sea and the underexplored  
567 "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of*  
568 *America* 103(32), 12115-12120.

569 Pedros-Alio, C. (2012) The rare bacterial biosphere. *Ann Rev Mar Sci* 4, 449-466.

570 Yu, K. and Zhang, T. (2012) Metagenomic and metatranscriptomic analysis of microbial  
571 community structure and gene expression of activated sludge. *PLOS one* 7(5), e38183.

572 Kim, T.S., Jeong, J.Y., Wells, G.F. and Park, H.D. (2013) General and rare bacterial taxa  
573 demonstrating different temporal dynamic patterns in an activated sludge bioreactor.  
574 *Applied Microbiology and Biotechnology* 97(4), 1755-1765.

575 Greenberg, A.E., Clesceri, L.S. and Eaton, A.D., eds. (1992) *Standard methods for the*  
576 *examination of water and wastewater*, American Public Health Association Publications,  
577 Washington, D.C., U.S.A.

578 Courtens, E.N.P., Vlaeminck, S.E., Vilchez-Vargas, R., Verliefde, A., Jauregui, R., Pieper,  
579 D.H. and Boon, N. (2014) Trade-off between mesophilic and thermophilic denitrification:  
580 Rates vs. sludge production, settleability and stability. *Water Research* 63, 234-244.

581 Judd, S. and Judd, C., (eds.) (2006) *The MBR Book: Principles and Applications of*  
582 *Membrane Bioreactors in Water and Wastewater Treatment*, Elsevier Science, Oxford,  
583 UK.

584 Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) Protein measurement  
585 with the Folin phenol reagent. *Journal of Biological Chemistry* 193(1), 265-275.

586 Vilchez-Vargas, R., Geffers, R., Suárez-Diez, M., Conte, I., Waliczek, A., Kaser, V.S.,  
587 Kralova, M., Junca, H. and Pieper, D.H. (2013) Analysis of the microbial gene landscape  
588 and transcriptome for aromatic pollutants and alkane degradation using a novel internally  
589 calibrated microarray system. *Environ Microbiol* 15(4), 1016-1039.

590 Bohorquez, L.C., Delgado-Serrano, L., López, G., Osorio-Forero, C., Klepac-Ceraj, V.,  
591 Kolter, R., Junca, H., Baena, S. and Zambrano, M.M. (2012) In-depth characterization via  
592 complementing culture-independent approaches of the microbial community in an acidic  
593 hot spring of the Colombian Andes. *Microbial Ecology* 63(1), 103-115.

594 Camarinha-Silva, A., Jáuregui, R., Chaves-Moreno, D., Oxley, A.P.A., Schaumburg, F.,  
595 Becker, K., Wos-Oxley, M.L. and Pieper, D.H. (2014) Comparing the anterior rare



596 bacterial community of two discrete human populations using Illumina amplicon  
597 sequencing. *Environ Microbiol* 16(9), 2939-2952.

598 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B.,  
599 Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B.,  
600 Thallinger, G.G., Van Horn, D.J. and Weber, C.F. (2009) Introducing mothur: open-  
601 source, platform-independent, community-supported software for describing and  
602 comparing microbial communities. *Applied and environmental microbiology* 75(23), 7537-  
603 7541.

604 McMurdie, P.J. and Holmes, S. (2013) Phyloseq: An R package for reproducible  
605 interactive analysis and graphics of microbiome census data. *PLOS one* 8(4), e61217.

606 Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B.,  
607 Simpson, G.L., Solymos, P., Stevens, M.H.H. and Wagner, H. (2013) Vegan: Community  
608 ecology package. *R package version 2.0-9*.

609 Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-  
610 Alfaro, A., Kuske, C.R. and Tiedje, J.M. (2014) Ribosomal Database Project: Data and  
611 tools for high throughput rRNA analysis. *Nucleic Acids Res* 42(Database issue), D633-  
612 642.

613 Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R. (2007) Naïve Bayesian classifier for  
614 rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and  
615 environmental microbiology* 73(16), 5261-5267.

616 Ramette, A. (2007) Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol*  
617 62(2), 142-160.

618 Harrell, F.E., Jr (2014) Hmisc: Harrell Miscellaneous. *R package version 3.14-6*.

619 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N.,  
620 Schwikowski, B. and Ideker, T. (2003) Cytoscape: A software environment for integrated  
621 models of biomolecular interaction networks. *Genome Res* 13(11), 2498-2504.

622 Gonzalez-Martinez, A., Rodriguez-Sanchez, A., Lotti, T., Garcia-Ruiz, M.-J., Osorio, F.,  
623 Gonzalez-Lopez, J. and van Loosdrecht, M.C.M. (2016) Comparison of bacterial  
624 communities of conventional and A-stage activated sludge systems. *Scientific Reports* 6,  
625 18786.

626 Saikaly, P.E. and Oerther, D.B. (2004) Bacterial competition in activated sludge:  
627 Theoretical analysis of varying solids retention times on diversity. *Microbial Ecology*  
628 48(2), 274-284.

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

632 Saikaly, P.E., Stroot, P.G. and Oerther, D.B. (2005) Use of 16S rRNA gene terminal  
633 restriction fragment analysis to assess the impact of solids retention time on the bacterial  
634 diversity of activated sludge. *Applied and environmental microbiology* 71(10), 5814-5822.

635 Bagchi, S., Tellez, B.G., Rao, H.A., Lamendella, R. and Saikaly, P.E. (2015) Diversity and  
636 dynamics of dominant and rare bacterial taxa in replicate sequencing batch reactors  
637 operated under different solids retention time. *Applied Microbiology and Biotechnology*  
638 99(5), 2361-2370.

639 Tan, T.W., Ng, H.Y. and Ong, S.L. (2008) Effect of mean cell residence time on the  
640 performance and microbial diversity of pre-denitrification submerged membrane  
641 bioreactors. *Chemosphere* 70(3), 387-396.

642 Teksoy Başaran, S., Aysel, M., Kurt, H., Ergal, İ., Akarsubaşı, A., Yağcı, N., Doğruel, S.,  
643 Çokgör, E.U., Keskinler, B., Sözen, S. and Orhon, D. (2014) Kinetic characterization of  
644 acetate utilization and response of microbial population in super fast membrane bioreactor.

645 Journal of Membrane Science 455(0), 392-404.

646 van der Gast, C.J., Jefferson, B., Reid, E., Robinson, T., Bailey, M.J., Judd, S.J. and

647 Thompson, I.P. (2006) Bacterial diversity is determined by volume in membrane

648 bioreactors. *Environ Microbiol* 8(6), 1048-1055.

649 van der Gast, C.J., Ager, D. and Lilley, A.K. (2008) Temporal scaling of bacterial taxa is

650 influenced by both stochastic and deterministic ecological factors. *Environ Microbiol*

651 10(6), 1411-1418.

652 Winkler, S., Matsche, N., Gamperer, T. and Dum, M. (2004) Sewage-treatment under

653 substantial load variations in winter tourism areas - A full case study. *Water Science and*

654 *Technology* 50(7), 147-155.

655 Curtis, T.P., Head, I.M. and Graham, D.W. (2003) Peer reviewed: Theoretical ecology for

656 engineering biology. *Environmental Science & Technology* 37(3), 64A-70A.

657 Preston, F.W. (1960) Time and space and the variation of species. *Ecology* 41(4), 612-627.

658 Wells, G.F., Park, H.D., Eggleston, B., Francis, C.A. and Criddle, C.S. (2011) Fine-scale

659 bacterial community dynamics and the taxa-time relationship within a full-scale activated

660 sludge bioreactor. *Water Research* 45(17), 5476-5488.

661 Shade, A., Caporaso, J.G., Handelsman, J., Knight, R. and Fierer, N. (2013) A meta-

662 analysis of changes in bacterial and archaeal communities with time. *The ISME Journal*

663 7(8), 1493-1506.

664 Hai, R., Wang, Y., Wang, X., Li, Y. and Du, Z. (2014) Bacterial community dynamics and

665 taxa-time relationships within two activated sludge bioreactors. *PLOS one* 9(3), 8 pp.

666 Ibarbalz, F.M., Perez, M.V., Figuerola, E.L. and Erijman, L. (2014) The bias associated

667 with amplicon sequencing does not affect the quantitative assessment of bacterial

668 community dynamics. *PLOS one* 9(6), e99722.

669 Lynch, M.D.J. and Neufeld, J.D. (2015) Ecology and exploration of the rare biosphere.

670 *Nature Reviews Microbiology* 13(4), 217-229.

671 Campbell, B.J., Yu, L., Heidelberg, J.F. and Kirchman, D.L. (2011) Activity of abundant

672 and rare bacteria in a coastal ocean. *Proceedings of the National Academy of Sciences of*

673 *the United States of America* 108(31), 12776-12781.

674 Gobet, A., Quince, C. and Ramette, A. (2010) Multivariate cutoff level analysis

675 (MultiCoLA) of large community data sets. *Nucleic Acids Res* 38(15), e155.

676 Ju, F., Xia, Y., Guo, F., Wang, Z. and Zhang, T. (2014) Taxonomic relatedness shapes

677 bacterial assembly in activated sludge of globally distributed wastewater treatment plants.

678 *Environ Microbiol* 16(8), 2421-2432.

679 Shade, A., Jones, S.E., Caporaso, J.G., Handelsman, J., Knight, R., Fierer, N. and Gilbert,

680 J.A. (2014) Conditionally rare taxa disproportionately contribute to temporal changes in

681 microbial diversity. *mBio* 5(4).

682 Saunders, A.M., Albertsen, M., Vollertsen, J. and Nielsen, P.H. (2016) The activated

683 sludge ecosystem contains a core community of abundant organisms. *The ISME Journal*

684 10(1), 11-20.

685 Magurran, A.E. and Henderson, P.A. (2003) Explaining the excess of rare species in

686 natural species abundance distributions. *Nature* 422(6933), 714-716.

687 Ulrich, W. and Zalewski, M. (2006) Abundance and co-occurrence patterns of core and

688 satellite species of ground beetles on small lake islands. *Oikos* 114(2), 338-348.

689 Barberan, A., Bates, S.T., Casamayor, E.O. and Fierer, N. (2012) Using network analysis

690 to explore co-occurrence patterns in soil microbial communities. *The ISME Journal* 6(2),

691 343-351.

692 Berry, D. and Widder, S. (2014) Deciphering microbial interactions and detecting keystone

693 species with co-occurrence networks. *Front Microbiol* 5, 219.

694 Paine, R.T. (1995) A Conversation on Refining the Concept of Keystone Species.  
695 Conservation Biology 9(4), 962-964.

696 McIlroy, S.J., Saunders, A.M., Albertsen, M., Nierychlo, M., McIlroy, B., Hansen, A.A.,  
697 Karst, S.M., Nielsen, J.L. and Nielsen, P.H. (2015) MiDAS: The field guide to the  
698 microbes of activated sludge, pp. Database, Vol. 2015, Article ID bav2062.

699 Hou, P.-b., Li, Y.-z., Wu, B.-h., Yan, Z.-c., Yan, B.-x. and Gao, P.-j. (2006) Cellulolytic  
700 complex exists in cellulolytic myxobacterium *Sorangium*. Enzyme and Microbial  
701 Technology 38(1–2), 273-278.

702 Rachid, S., Gerth, K., Kochems, I. and Müller, R. (2007) Deciphering regulatory  
703 mechanisms for secondary metabolite production in the myxobacterium *Sorangium*  
704 *cellulosum* So ce56. Molecular Microbiology 63(6), 1783-1796.

705 Kalmbach, S., Manz, W., Wecke, J. and Szewzyk, U. (1999) *Aquabacterium* gen. nov.,  
706 with description of *Aquabacterium citratiphilum* sp. nov., *Aquabacterium parvum* sp. nov.  
707 and *Aquabacterium commune* sp. nov., three in situ dominant bacterial species from the  
708 Berlin drinking water system. Int J Syst Bacteriol 49 Pt 2, 769-777.

709 Kojima, H. and Fukui, M. (2011) *Sulfuritalea hydrogenivorans* gen. nov., sp. nov., a  
710 facultative autotroph isolated from a freshwater lake. Int J Syst Evol Microbiol 61(Pt 7),  
711 1651-1655.

712 Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H. and Stackebrandt, E. (Eds.),  
713 (2006) The Prokaryotes: A Handbook on the Biology of Bacteria., Springer-Verlag, New  
714 York.

715

**Table 1**[Click here to download Table: Table 1.docx](#)

**Table 1:** Average values of environmental and operational variables throughout the study period, with standard deviations. Averages that differ by more than a factor two between the high- and low-rate system are indicated in bold. N = number of data points. The *p*-values indicate the significance level of pairwise comparisons between the high-rate and low-rate values.

Environmental variables	Abbreviation	High-rate		Low-rate		n	<i>p</i> -value	
Day of sampling	Time	Day 0 (Oct 2013) to 273 (Jul 2014)		d		38		
Temperature	Temperature	10.1 (min) - 20.4 (max)		°C		38		
Rainfall	Rainfall	0 (min) - 13.4 (max)		mm/d		38		
Recirculation factor of final effluent back to influent	R.factor	1.3 ± 0.7		fraction		38		
<b>BOD concentration of influent</b>	<b>BOD</b>	<b>100.4</b> ±	<b>29.2</b>	<b>47.7</b> ±	<b>12.6</b>	<b>mg L<sup>-1</sup></b>	<b>37</b>	<b>1.58 x 10<sup>-13</sup></b>
<b>Floc size (volume-weighted average diameter)</b>	<b>D<sub>4,3</sub></b>	<b>256.7</b> ±	<b>83.6</b>	<b>87.1</b> ±	<b>8.3</b>	<b>µm</b>	<b>27</b>	<b>3.38 x 10<sup>-11</sup></b>
<b>Hydraulic retention time (nominal)</b>	<b>HRT<sub>nom</sub></b>	<b>0.024</b> ±	<b>0.012</b>	<b>0.188</b> ±	<b>0.102</b>	<b>d</b>	<b>38</b>	<b>5.57 x 10<sup>-12</sup></b>
<b>Sludge retention time of reactor + settling system</b>	<b>SRT<sub>svst</sub></b>	<b>1.74</b> ±	<b>0.53</b>	<b>34.4</b> ±	<b>28.8</b>	<b>d</b>	<b>37</b>	<b>4.49 x 10<sup>-8</sup></b>
<b>Sludge-specific loading rate</b>	<b>SLR</b>	<b>2.13</b> ±	<b>0.67</b>	<b>0.11</b> ±	<b>0.03</b>	<b>g BOD g<sup>-1</sup> VSS d<sup>-1</sup></b>	<b>37</b>	<b>8.33 x 10<sup>-20</sup></b>
COD removal efficiency	COD.removed	0.54 ±	0.11	0.70 ±	0.07	fraction	37	2.22 x 10 <sup>-10</sup>
COD/N ratio of influent	COD/N	10.8 ±	2.1	6.3 ±	1.3	mg mg <sup>-1</sup>	37	2.59 x 10 <sup>-16</sup>
Kjeldahl nitrogen concentration of influent	KjN	23.2 ±	4.7	20.8 ±	3.8	mg L <sup>-1</sup>	37	1.58 x 10 <sup>-2</sup>
Nitrogen removal efficiency	N.removed	0.33 ±	0.11	0.52 ±	0.12	fraction	37	3.48 x 10 <sup>-10</sup>
Observed sludge growth yield	Y <sub>obs</sub>	0.67 ±	0.23	0.50 ±	0.54	g TSS g <sup>-1</sup> COD	37	n.s.
Phosphorus concentration (incoming)	P	4.3 ±	1.1	3.0 ±	1.2	mg L <sup>-1</sup>	37	3.14 x 10 <sup>-6</sup>
Phosphorus removal efficiency	P.removed	0.47 ±	0.18	0.45 ±	0.16	fraction	37	n.s.
Proteinaceous extracellular polymeric substances	EPS.P	37.6 ±	8.2	73.9 ±	26.3	mg BSA g <sup>-1</sup> VSS	17	3.49 x 10 <sup>-5</sup>
Sludge volume index	SVI	76.5 ±	14.3	120.1 ±	15.9	mL g <sup>-1</sup>	38	4.06 x 10 <sup>-20</sup>
TSS concentration	TSS	2780 ±	545	3371 ±	444	mg L <sup>-1</sup>	38	1.62 x 10 <sup>-6</sup>
VSS/TSS ratio in high-rate system	VSS.TSS	0.79 ±	0.04	n.a.		fraction	38	
Oxygen concentration in second compartment high-rate	O <sub>2</sub> .A2	0.44 ±	0.21	n.a.		mg L <sup>-1</sup>	38	
Oxygen concentration in third compartment high-rate	O <sub>2</sub> .A3	1.74 ±	0.55	n.a.		mg L <sup>-1</sup>	38	

**Table 2:** Variation partitioning using canonical correspondence analysis (CCA) on the total community of the high-rate and low-rate system, and of the three sub-communities. For each CCA analysis, only those environmental variables were included that correlated significantly to the ordination axes of an unconstrained correspondence analysis.

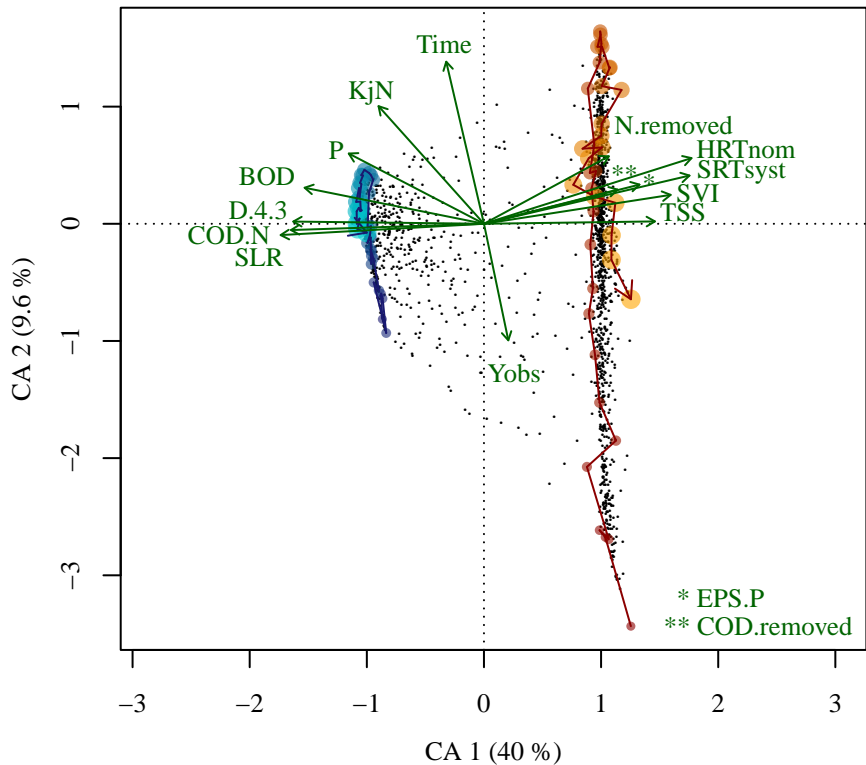
	<b>High-rate system</b>		<b>Low-rate system</b>	
	<b>Significant variables</b>	<b>% of variation</b>	<b>Significant variables</b>	<b>% of variation</b>
<b>Total</b>	Time, HRT <sub>nom</sub> , KjN, P, Temperature, Y <sub>obs</sub>	47.5%	BOD, D <sub>4,3</sub> , Time, HRT <sub>nom</sub> , KjN, SVI, Temperature	55.9%
<b>Continuously abundant</b>	Time, HRT <sub>nom</sub> , KjN, P, Temperature	45.1%	BOD, D <sub>4,3</sub> , Time, HRT <sub>nom</sub> , KjN, N.removed, Temperature	60.6%
<b>Transitional</b>	Time, HRT <sub>nom</sub> , KjN, P, Temperature, Y <sub>obs</sub>	51.0%	BOD, D <sub>4,3</sub> , Time, HRT <sub>nom</sub> , KjN, SVI, Temperature	60.1%
<b>Continuously rare</b>	Time, HRT <sub>nom</sub> , KjN, Temperature, Y <sub>obs</sub>	28.5%	BOD, D <sub>4,3</sub> , Time, HRT <sub>nom</sub> , KjN, SVI, Temperature	44.4%

**Table 3**[Click here to download Table: Table 3.docx](#)

**Table 3:** Distribution of phylotypes and sequences of the continuously abundant, transitional and continuously rare sub-communities over the entire time series (38 samples) of the high-rate and low-rate communities. At each time point, abundant and rare phylotypes were distinguished by a 0.1% relative abundance threshold.

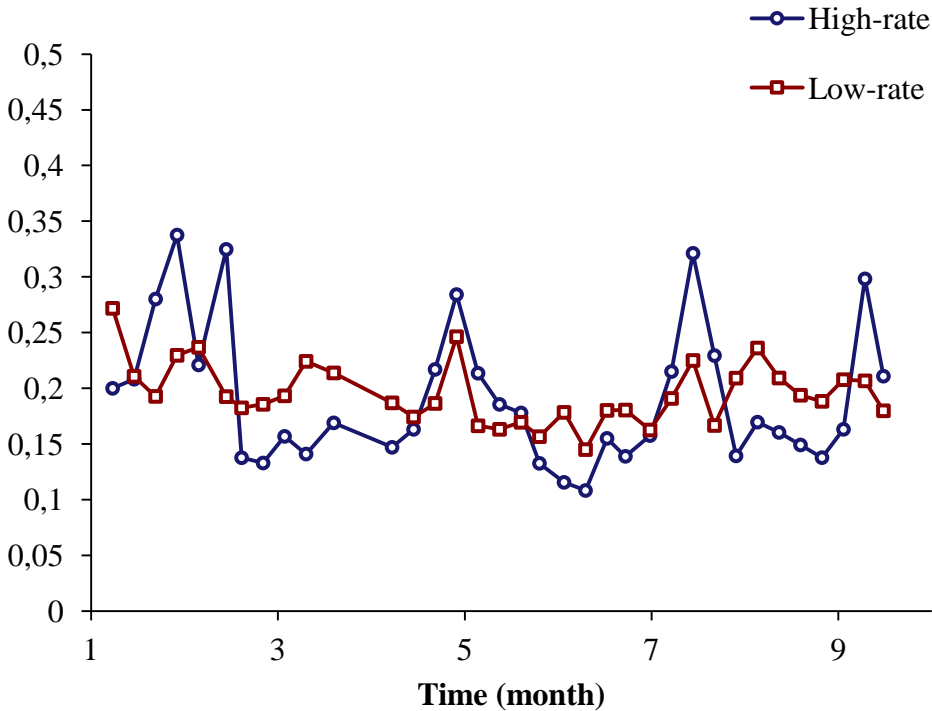
	High-rate				Low-rate			
	Phylotypes		Sequences		Phylotypes		Sequences	
Continuously abundant	16	<b>2.1%</b>	$3.8 \times 10^5$	<b>65.2%</b>	34	<b>2.3%</b>	$2.5 \times 10^5$	<b>43.2%</b>
Transitional	237	<b>30.5%</b>	$1.8 \times 10^5$	<b>31.5%</b>	547	<b>36.5%</b>	$2.8 \times 10^5$	<b>49.4%</b>
Continuously rare	523	<b>67.4%</b>	$1.9 \times 10^4$	<b>3.3%</b>	919	<b>61.3%</b>	$4.3 \times 10^4$	<b>7.4%</b>
Total	776		$5.8 \times 10^5$		1500		$5.8 \times 10^5$	

Figure 1 with caption 273



**Figure 1:** Correspondence analysis (CA) of the combined high-rate (blue) and low-rate (red) communities from October 2013 to July 2014. Phylotypes are shown as dots. Samples are shown as circles with increasing size in chronological order, and connected by a blue or red arrow. Environmental variables that significantly correlate to the ordination are plotted as green arrows. Abbreviations are the same as in **Table 1**. Percentages indicate the relative contribution of each axis to total inertia.

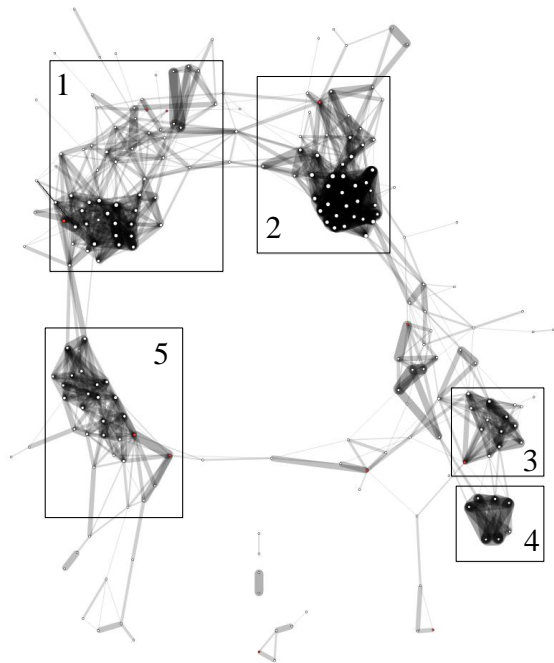
Figure 2 with caption  
Bray-Curtis dissimilarity



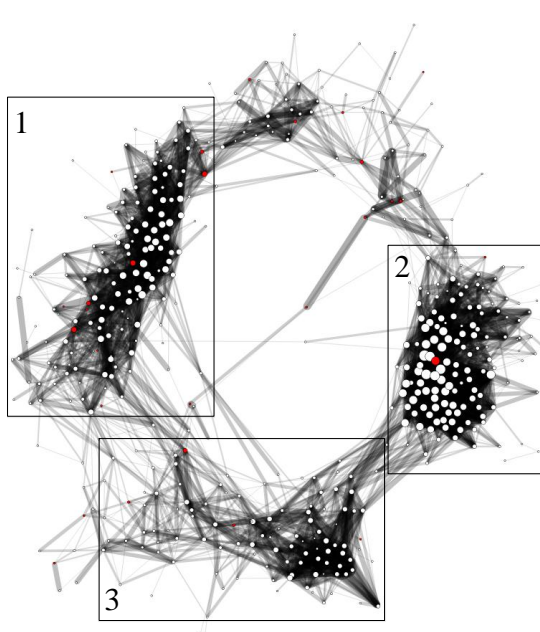
**Figure 2:** Moving-window analysis of the Bray-Curtis dissimilarity between samples with a one-week interval, for the high-rate (blue) and low-rate (red) communities.



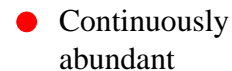
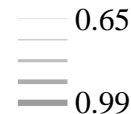
**Figure 3 with caption**  
High-rate system



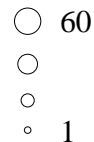
Low-rate system



Pearson coefficient

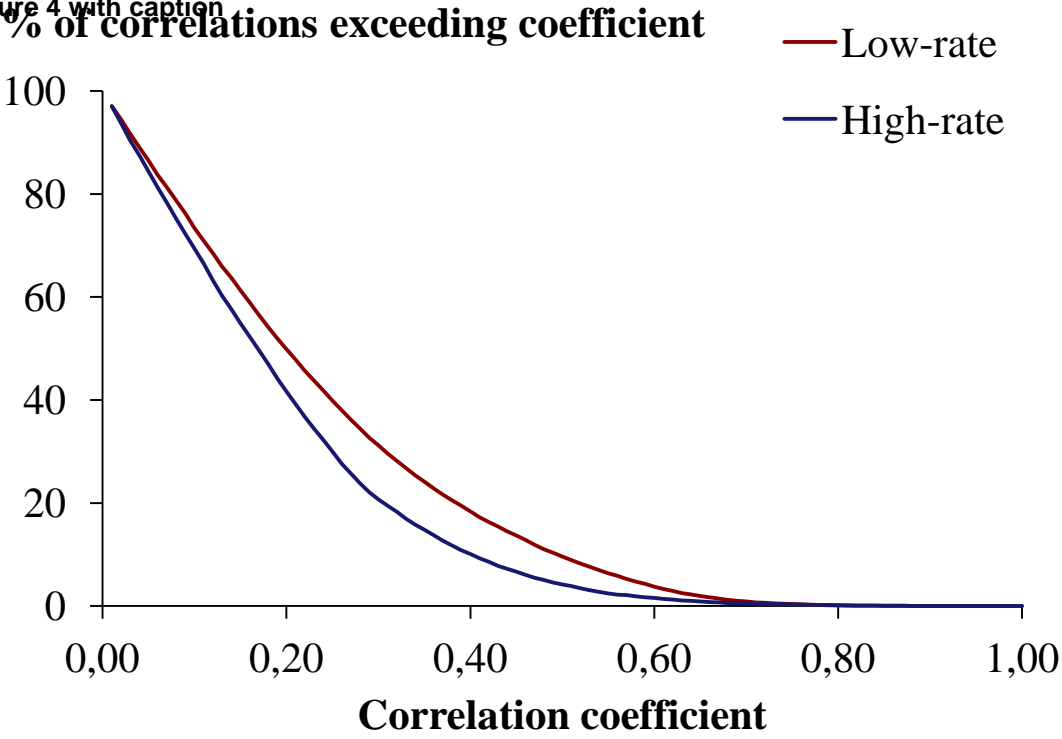


Node degree



**Figure 3:** Co-occurrence network of the high-rate (left) and low-rate (right) communities, based on Pearson correlations. Positive correlations ( $r > 0.65$ ,  $p < 0.05$ ) were considered for the continuously abundant (red) and transitional (grey) sub-communities; continuously rare phylotypes were excluded from the analysis. Singleton nodes (i.e., nodes not connected to any other node) are not visualized. The node size represents the node degree, and the line thickness represents the strength of the correlation. Rectangles indicate different clusters within each network, as visually identified.

Figure 4 with caption



**Figure 4:** Percentage of all Spearman correlations between individual phylotypes and environmental variables for which the absolute coefficient exceeds a given value between  $r = 0$  and  $r = 1$ , in the high-rate and low-rate community.