

1 **Title:** Neighbour urban wastewater treatment plants display distinct profiles of bacterial
2 community and antibiotic resistance genes

3 **Authors:** Telma Fernandes^{a,1}, Ivone Vaz-Moreira^a, Célia M. Manaia^{a,*}

4 **Affiliation:** ^a Universidade Católica Portuguesa, CBQF - Centro de Biotecnologia e Química
5 Fina – Laboratório Associado, Escola Superior de Biotecnologia, Rua Arquiteto Lobão Vital,
6 172, 4200-374 Porto, Portugal

7 Current address: ¹ Telma Fernandes current affiliation is Instituto de Tecnologia Química e
8 Biológica António Xavier, Universidade Nova de Lisboa (ITQB NOVA), Oeiras 2780-157,
9 Portugal

10 **Corresponding author (*):**

11 C.M. Manaia, PhD
12 Escola Superior de Biotecnologia, Universidade Católica Portuguesa
13 Rua Arquiteto Lobão Vital, 172
14 4200-374 Porto; Portugal
15 Tel: +351 22 5580059, Fax: +351 22 5090351
16 e-mail: cmanaia@porto.ucp.pt

17

18

19 **ACKNOWLEDGMENTS:**

20 The authors gratefully acknowledge the support of the staff of the UWTPs and supplying entities
21 that made this study possible by providing the water samples; Christophe Merlin that kindly
22 provided the integrase gene qPCR protocol and Gonçalo Macedo and Jaqueline Rocha for
23 technical assistance on sampling, DNA extraction, and qPCR analysis.

24 This work was funded by National Funds from FCT – Fundação para a Ciência e a Tecnologia
25 through project WaterJPI/0001/2013 STARE – “Stopping Antibiotic Resistance Evolution” and
26 UID/Multi/50016/2013. IVM was supported by the FCT grant (SFRH/BPD/87360/2012).

27

28

29 **ABSTRACT**

30 Urban wastewater treatment plants (UWTPs) are among the major recipients of antibiotic-
31 resistant bacteria (ARB), antibiotic resistance genes (ARGs) and antibiotic residues in urban
32 environments. Although during treatment bacteria of human and animal origin are removed,
33 some are able to survive, persisting in the final effluent. The occurrence of these bacteria,
34 especially those harbouring ARGs, may have a direct impact on the quality of the treated
35 wastewater that is returned to the environment. In this study, we aimed to assess if the final
36 effluent bacterial communities of three UWTPs (PT1, PT2, and PT3) located next to each other,
37 were distinct and if such differences were related with the antibiotic resistance profiles.
38 It was observed that the bacterial community (16S rRNA gene Illumina sequencing) and load
39 of selected ARGs of final effluent differed among the three UWTP, irrespective of sampling
40 time. Members of the families *Aeromonadaceae*, *Campylobacteraceae*, *Veillonellaceae*,
41 *Weeksellaceae*, and *Porphyromonadaceae* were observed to be positively correlated with some
42 ARGs (*bla_{CTX-M}*, *bla_{OXA-A}*, *bla_{SHV}*) and *intl1* ($p < 0.05$), while *Intrasporangiaceae* were observed to
43 be negatively correlated. While *Aeromonadaceae* are recognized relevant ARGs harbours, the
44 other bacterial families may represent bacteria that co-exist with the ARGs hosts, which may
45 belong to minor bacterial groups, omitted in the analyses. These findings suggest the
46 importance of bacterial dynamics during treatment to the ARB&ARGs removal, a rationale that
47 may contribute to design new strategies to apply in the UWTPs to prevent the spread of
48 antibiotic resistance.

49

50 **KEYWORDS:** final effluent; selection; bacterial diversity; antibiotic resistance genes; correlation
51 analyses

52 1. INTRODUCTION

53 Urban wastewater treatment plants (UWTPs) have received much attention as major recipients
54 of antibiotic-resistant bacteria (ARB), antibiotic resistance genes (ARGs) (Rizzo et al. 2013) and
55 antibiotic residues (Michael et al. 2013) in urban environments. Their capacity to remove ARB
56 and ARGs has also been largely discussed and demonstrated (Manaia et al. 2016; Bengtsson-
57 Palme et al. 2016; Wu et al. 2018). Nevertheless, most authors agree that ARB and ARGs
58 loads released by well-functioning UWTPs may still have a negative impact on the environment,
59 with serious implications to the human health (Berendonk et al. 2015). While it is consensual
60 that it is important to improve wastewater treatment processes in order to maximize the removal
61 of ARB and ARGs, not much is known about the best strategies to achieve such a goal
62 (Berendonk et al. 2015; Manaia et al. 2016; Vikesland et al. 2017). Although wastewater
63 disinfection may promote the reduction of the total microbial loads and of the bacterial diversity,
64 it can create the perfect conditions for the bacterial regrowth of certain groups of bacteria
65 (Manaia et al. 2016; Becerra-Castro et al. 2017; Moreira et al. 2018).

66 The maintenance or propagation of antibiotic resistance during wastewater treatment may be
67 due to two major driving forces, ARG horizontal gene transfer and/or ARB survival or selection
68 (Berendonk et al. 2015; Vikesland et al. 2017). Despite the importance that horizontal gene
69 transfer may have on ARGs spread, it may be not the most relevant factor to dictate the fate of
70 antibiotic resistance during wastewater treatment (Figueira et al. 2011a; Figueira et al. 2011b;
71 Manaia et al. 2016). The survival or not of specific bacterial lineages during treatment may have
72 consequences on the abundance of ARGs in the final effluent. In this study, we aimed to
73 investigate if ARGs released in the final effluent of a UWTP are related with the diversity of
74 bacteria present in that effluent, at least with some of the groups surviving the wastewater
75 treatment, comprising bacteria frequently associated with humans and animals. In this rationale,
76 it is assumed that the understanding of which bacterial groups are the major vehicles or the
77 ecologic surrogates of ARGs could be a good strategy to identify suitable bacterial indicators of
78 ARGs that survive treatment. This identification could be the basis to manage wastewater
79 treatment processes in order to maximize the reduction or eliminate such bacterial groups. To
80 test the aforementioned hypothesis, we compared the final effluent of three UWTPs, located in
81 the same region (Northern Portugal), and explored whether it was possible to find significant

82 correlations between the bacterial community composition and the profile of a selected group of
83 ARGs.

84

85 **2. MATERIALS AND METHODS**

86 **2.1. Sampling**

87 Treated wastewater samples were collected in three UWTPs (PT1, PT2, and PT3) located,
88 within a region of approximately 7 km, in Northern Portugal. These UWTPs, serve populations
89 of 150 000 - 170 000 inhabitants equivalent, receive a daily flow of approximately 22 000 m³
90 (PT1), 33 000 m³ (PT2), and 24 000 m³ (PT3), and all have primary and secondary treatments
91 operating with conventional activated sludge (CAS). PT2 is the only one reporting the reception
92 of hospital effluents and PT3 is the only one with UV disinfection of the final effluent. No
93 information was provided regarding the reception of hospital effluents by PT1 or PT3. Soon after
94 this study, PT1 was temporarily closed for improvement restoration.

95 Composite samples (24 h) were collected from PT1 and PT2 in four campaigns, in early Spring
96 (average temperature of 9 °C) and early Autumn (average temperature of 17 °C), specifically in
97 three consecutive days (Tuesday - Thursday) in March 2015 (M_2015), October 2015
98 (O_2015), March 2016 (M_2016), and September 2016 (S_2016). Due to technical problems at
99 PT1, it was not possible to sample this UWTP in March 2016. Grab samples of PT3 were
100 collected in June 2015 (J_2015), July 2015 (JL_2015), September 2015 (S_2015), and
101 March 2016 (M_2016). Samples were stored in ice and transported to the laboratory to be
102 processed within 12 h.

103 **2.2. DNA extraction**

104 Volumes of 100 to 250 mL of final effluent were filtered, in triplicate, through sterile
105 polycarbonate membranes (0.22 µm porosity, Whatman, UK) that were stored at -80 °C. DNA
106 was extracted from each of the three replicates with the PowerWater® DNA Isolation Kit
107 (MOBIO Laboratories Inc., CA, USA) according to the manufacturer's instructions. DNA
108 extract's concentration was measured using the Qubit 3.0 Fluorometer (ThermoFisher

109 Scientific, USA). A total of 33 DNA extracts, corresponding to 7 independent sampling dates,
110 were stored at -20 °C until further analysis.

111 **2.3. Bacterial community composition**

112 The bacterial community composition was analysed in the 33 DNA extracts, based on
113 MiSeq® Illumina DNA sequencing (Genoinseq, Cantanhede, Portugal) of the V3 - V4
114 hypervariable region of the 16S rRNA gene using the
115 Bakt_341F 5' - CCTACGGGNGGCWGCAG - 3' and
116 Bakt_805R 5' - GACTACHVGGGTATCTAATCC - 3' primers set, according to manufacturer's
117 instructions (Illumina, San Diego, CA, USA). A filtering based on the reads size and quality
118 trimming was performed using the software PRINSEQ (Schmieder and Edwards 2011). Reads
119 shorter than 200 bp and with average quality scores lower than 25 were excluded. Sequence
120 reads were demultiplexed automatically by the Illumina® Miseq® sequencer using the CASAVA
121 package (Illumina, San Diego, CA, USA) and paired-end reads were merged using a QIIME
122 script. Chimeric sequences were removed using USEARCH v6.1 (Edgar 2010). After the quality
123 control, one of the three replicates of PT3 (September 2015) was removed from the analyses
124 due to the low number of initial reads (n=37 932). Sequences with nucleotide identity higher
125 than 97% were assigned to Operational Taxonomic Units (OTUs) (Edgar 2010) and to
126 taxonomic groups using the Greengenes Database version 13_8 (updated: August 2013)
127 (DeSantis et al. 2006). Singletons, chimeras as well as OTUs assigned to non-bacterial groups,
128 such as chloroplasts or mitochondria, were removed from the data set. After this procedure
129 3 109 382 good-quality sequences were obtained, grouped into 32 084 OTUs. The downstream
130 analyses were performed after normalization with the cumulative sum scaling (CSS) (Paulson et
131 al. 2014) of the data (except the computation of the alpha diversity metrics). The alpha diversity
132 was evaluated based on bacterial richness (number of OTUs), diversity indices Shannon
133 (Shannon and Weaver 1963), phylogenetic diversity (PD) whole tree (Faith 1992), and Simpson
134 (Simpson 1949), and the richness estimator Chao1 (Chao 1984), using the QIIME 1 pipeline
135 (Caporaso et al. 2010). Alpha diversity was represented as the mean value for each index and
136 the respective standard deviation (SD) values and differences among UWTP effluents
137 calculated by analysis of variance (ANOVA) with post - hoc Turkey's test ($p < 0.01$). Alpha

138 rarefaction plot of observed OTUs was performed to verify the sequence coverage. All the plots
139 reached a plateau phase meaning that the final data is a good representation of the bacterial
140 diversity of the wastewater effluent. Beta diversity indices were assessed using the weighted
141 UniFrac metric distances (Lozupone and Knight 2005) from QIIME pipeline (Caporaso et al.
142 2010) and the results presented as principal coordinates analysis (PCoA). This analysis was
143 also performed by taking into consideration the UWTPs individually in order to assess possible
144 seasonal variations. The relative abundance of the bacterial community composition at different
145 taxonomic levels was compared based on the ANOVA and Tukey-Kramer post-hoc tests
146 ($p < 0.01$) and the p - values were corrected for multiple testing using the Benjamini - Hochberg
147 false discovery rate (FDR) (Benjamini and Hochberg 1995), using STAMP v2.1.3 software
148 (Parks et al. 2014). Pairs of samples with an FDR value < 0.01 were considered significantly
149 different.

150 The sequence data files that support the findings of this study have been deposited in GenBank
151 within the BioProject PRJNA478220, Sequence Read Archive (SRA) SRS3472441, with the
152 accession numbers SRX4313251-SRX4313282.

153 **2.4. Quantification of antibiotic resistance genes and class 1 integron-integrase**

154 Quantitative polymerase chain reaction (qPCR, StepOne™ Real-Time PCR System, Life
155 Technologies, Carlsbad, CA, USA) was used to measure the abundance (per mL of sample) of
156 seven ARGs (*bla*_{CTX-M}, *bla*_{OXA-A}, *bla*_{SHV}, *bla*_{TEM}, *sul1*, *sul2*, and *qnrS*) and class 1
157 integron - integrase (*int1*). The 16S rRNA gene was also quantified, as a measure of the total
158 bacterial load, and used to determine ARGs and *int1* prevalence values (per 16S rRNA gene
159 copy number). Calibration curves, built based on adequate dilutions of standards of each
160 analysed gene, were produced in parallel with the test samples. Each DNA extract was tested in
161 duplicate for each run. The list of primers and protocols used, based on SYBR Green detection,
162 have been described before (Narciso-da-Rocha et al. 2018). One-way analysis of variance
163 (ANOVA) and Tukey's post-hoc tests were used to assess statistically significant differences
164 ($p < 0.01$) of prevalence and/or abundance of 16S rRNA gene, ARGs or *int1* using
165 GraphPad Prism 6.00 software (GraphPad Software, San Diego, CA). Principal component

166 analysis (PCA) based on the prevalence of ARGs was performed using Canoco 5.01 software
167 (Leps and Smilauer 2014).

168 **2.5. Comparative analyses of bacterial community composition and ARGs and *intl1***

169 A Redundancy Analysis (RDA) (Wollenberg 1977) was performed to explore possible
170 relationships between the bacterial community structure (families with relative abundance > 1%)
171 and the ARGs and *intl1* prevalence (target gene copy number / 16S rRNA gene copy number)
172 for all the UWTPs, using Canoco 5.01 software (Leps and Smilauer 2014). The significance of
173 the environmental variables (ARGs and *intl1*) was tested using Monte Carlo permutation tests
174 (999 unrestricted permutations, $p < 0.05$ and $p < 0.01$) after FDR adjustment. The same data
175 set was used to assess Pearson's correlations, testing each UWTP individually, to infer about
176 possible relationships between the bacterial community composition and genes. Pearson's
177 correlation was performed using GraphPad Prism 6.00 software (GraphPad Software, San
178 Diego, CA).

179

180 **3. RESULTS**

181 **3.1. Bacterial community**

182 The analysis of the V3 - V4 region of the 16S rRNA gene led to the identification of 32 084
183 OTUs in the final effluent of the three UWTPs. PT1 showed the highest number of OTUs
184 ($7\,238 \pm 861$), followed by PT2 ($6\,732 \pm 1\,114$) and PT3 ($5\,704 \pm 996$). These values were in
185 agreement with the observation that PT3 displayed the lowest Chao1 richness estimator value
186 ($p < 0.01$). However, this difference was not expressed in the Shannon's and Simpson's
187 diversity indices, significantly higher in PT3 and PT2 than in PT1 ($p < 0.01$) (Table 1). The
188 whole tree phylogenetic diversity index (PD), which measure the shared phylogenetic distance
189 among taxa in each sample, did not reveal significant differences of the bacterial communities
190 ($p > 0.01$) in the final effluents of the three UWTPs.

191 The phyla *Proteobacteria* and *Bacteroidetes* predominated in the final effluents of all UWTPs
192 (PT1: 55.2 - 58.8% vs. 21.1 - 25.9%, PT2: 39.1 - 55.0% vs. 13.1 - 14.2%, and
193 PT3: 42.9 - 44.8% vs. 9.4 - 11.3%) (Fig 1). However, the candidate phyla, *TM7* (renamed as

194 Candidatus *Saccharibacteria*) and *OD1* (renamed as Candidatus *Parcubacteria*), made the
195 difference in PT2 and PT3, where they were also among the predominant, with relative
196 abundance values ranging 9.2 - 18.3% and 2.7 - 6.6%, respectively, for PT2, and 5.9 - 12.7%
197 and 8.4 - 14.0%, respectively, for PT3. These phyla were below 3.6% and 3.2%, respectively in
198 PT1. Predominant *Proteobacteria* classes were represented by OTUs affiliated to *Gamma*-
199 (PT1: 9.3 – 11.8%, PT2: 9.3 – 12.5%, and PT3: 7.1 – 10.9%) and *Betaproteobacteria*
200 (PT1: 31.2 – 38.4%, PT2: 16.1 – 26.6% and PT3: 13.6 – 19.7%), in particular of the order
201 *Burkholderiales* (PT1: 23.2 – 30.0%, PT2: 6.2 – 12.6 %, and PT3: 7.4 – 11.0%). *Bacteroidales*
202 was the predominant order of *Bacteroidetes* (PT1: 10.0 – 17.0%, PT2: 3.9 – 6.9% and
203 PT3: 2.5 – 6.7%). It was also noticeable that, at all taxonomic levels, unclassified OTUs were
204 more abundant in PT2 and PT3 than in PT1.

205 Based on the PCoA analysis, the bacterial community observed in the final effluent of the
206 three UWTPs formed different groups, which, according to this analysis, were not influenced
207 by sampling date (Fig. 2a). PT1 displayed the most variable bacterial community composition
208 (Fig. 2a), although PT2 and PT3 were also in separated groups (Fig. 2a). Given the fact that
209 OTUs were defined at a sequence identity value of 97%, it was surprisingly low the number
210 of OTUs shared by all UWTPs, of only 22.0% (Fig. 2b). This number was nevertheless higher
211 than that of unique OTUs, of 18.9%, 16.1%, and 10.8%, in PT1, PT2, and PT3, respectively.
212 Confirming the PCoA (Fig. 2a), PT3 and PT2 shared the highest number of OTUs (37.1%),
213 while PT1 shared 34.3% with PT2 and 26.9% with PT3 (Fig. 2b). In PT1, unique OTUs were
214 mainly *Proteobacteria* (58.4%) and *Bacteroidetes* (30.2%), of the families *Comamonadaceae*
215 and *Flavobacteriaceae*, together with bacteria of the lineage *GZKB119*. The remaining bacterial
216 phyla contributed less than 2.5% for the unique OTUs. In PT2 and PT3 unique OTUs were also
217 mainly *Proteobacteria* (PT2: 43.8%, PT3: 37.8%) and *Bacteroidetes* (PT2: 9.1%, PT3: 7.2%),
218 but other groups were observed, specifically *TM7* (Candidatus *Saccharibacteria*, PT2: 19.6%,
219 PT3: 7.4%) and *OD1* (PT2: 7.7%, PT3: 16.5%). PT3 had also a relevant percentage of unique
220 OTUs identified as *Planctomycetes* (> 5%). In PT2, unique OTUs identified at the family level
221 belonged to *Procabacteriaceae* and *Rhodocyclaceae*, while for PT3 no unique bacterial families
222 stood out.

223 A more in depth comparative analyses of the bacterial community composition in the final
224 effluent of the three UWTPs considered all families with relative abundance higher than 1%
225 (n=24 families). This comparison showed that PT1 differed from the other two UWTPs by a
226 higher relative abundance ($p < 0.01$) of [*Weeksellaceae*], *Flavobacteriaceae*, *GZKB119*,
227 *Porphyromonadaceae*, *Comamonadaceae*, and *Pseudomonadaceae*, and a lower relative
228 abundance ($p < 0.01$) of *Saprospiraceae* and *Bdellovibrionaceae* (Fig. 3). In PT2, the families
229 *Intrasporangiaceae*, *Chitinophagaceae* and *Procabacteriaceae* were more abundant than in the
230 other two UWTP ($p < 0.01$). PT3 differed from the others UWTPs on a higher relative
231 abundance of *Bdellovibrionaceae* and lower relative abundance of *Sphingomonadaceae* and
232 *Verrucomicrobiaceae* (Fig. 3). Curiously, PT2 and PT3 showed a higher abundance of
233 unclassified families (PT2: 52.6%, PT3: 68.7%) than PT1 (14.7%).

234 3.2. Antibiotic resistance genes

235 A set of ARGs (*bla*_{CTX-M}, *bla*_{TEM}, *bla*_{OXA-A}, *bla*_{SHV}, *sul1*, *sul2*, and *qnrS*) and class 1 integron-
236 integrase (*int1*) gene were analysed in terms of abundance and prevalence. The quantification
237 of the 16S rRNA gene is a measure of the bacteria content in all samples and supported the
238 estimation of prevalence values. Similar trends were observed for abundance and prevalence of
239 ARGs and *int1*, ranked as *bla*_{CTX-M} < *bla*_{SHV} < *bla*_{TEM} < *qnrS* < *sul2* < *bla*_{OXA-A} < *int1* < *sul1* (Fig. 4).
240 Comparing the different UWTPs, it was observed that PT1 presented significantly higher
241 ($p < 0.01$) abundance and prevalence of *bla*_{OXA-A} and *int1* and higher abundance of
242 16S rRNA gene, *bla*_{TEM}, and *qnrS* than PT2 and PT3. In PT3, the gene *bla*_{SHV} was more
243 abundant and *sul1* more prevalent than in the other plants. PT2 showed a higher prevalence of
244 *bla*_{SHV} and *sul2* ($p < 0.01$) than the other two UWTPs (Fig. 4a). Despite these significant
245 differences, the average abundance values of ARGs discharged by the different UWTPs
246 differed less than 1.2 log-unit for *bla*_{CTX-M}, *bla*_{SHV}, *bla*_{TEM}, *sul1*, *sul2*, *qnrS*, and *int1*, being less
247 than 2 log-units for the genes and *bla*_{OXA-A}. These results show that the average abundance
248 values of ARGs released by the three UWTPs analysed is very similar, independently of the
249 treatment used, suggesting that this may be associated with the region.

250 The bacterial community distribution seemed more stable than that of ARGs and *int1*
251 prevalence values. Even though, the general pattern of distribution was similar, with PT1
252 clustering apart from PT2 and PT3 (Fig. 4b and Fig. 2a). This observation motivated the search

253 for possible relationships between the bacterial community composition and structure and the
254 quantified genes in all UWTPs, using a Redundancy Analysis (RDA). According to this analysis,
255 the ARGs and *int1* prevalence could explain approximately 84% of the variation observed (Fig.
256 5). Among the genes showing significant correlation with the bacterial community distribution
257 were *bla*_{CTX-M}, *bla*_{OXA-A}, *bla*_{SHV}, and *int1* (Monte Carlo permutations, $p < 0.05$ for *bla*_{SHV} and
258 *int1*; $p < 0.01$ for *bla*_{CTX-M} and *bla*_{OXA-A}). The genes *bla*_{OXA-A} and *int1* were mainly associated
259 with PT1 samples and members of the families [*Weeksellaceae*], *Comamonadaceae*,
260 *Flavobacteriaceae*, *GZKB119*, and *Porphyromonadaceae* predominant in that UWTP. In
261 opposition, the genes *bla*_{CTX-M} and *bla*_{SHV} were mostly associated with PT2 and PT3 samples
262 and with members of the families *Bdellovibrionaceae*, *Gordoniaceae*, and *Saprospiraceae*
263 (Fig. 5 and Fig. 3).

264 For a further insight of the multivariate analysis provided by RDA, the comparative analyses
265 presented in Fig. 3 were revisited for inferring about possible statistically significant correlations
266 between specific bacterial families and ARGs and *int1* prevalence values. This analysis could
267 hint bacterial groups putatively associated with ARGs, either because they harbour some of
268 those genes or because they co-occur with the bacteria that harbour such genes (Ju et al.
269 2016). The identification of these bacteria is of interest, since being able to survive wastewater
270 treatment, they may contribute to the antibiotic resistance spread. The same analysis could also
271 suggest other groups whose presence might be associated with lower resistance prevalence, in
272 the case of negative correlations. The analysis revealed that members of the families
273 *Aeromonadaceae* and *Campylobacteraceae* (*Proteobacteria*), *Veillonellaceae* (*Firmicutes*), and
274 [*Weeksellaceae*] and *Porphyromonadaceae* (*Bacteroidetes*) were significantly positively
275 correlated with ARGs and *int1* in different UWTPs. None of these bacterial families was
276 correlated with the *sul1* gene. On the other hand, the *Intrasporangiaceae* were observed to be
277 negatively correlated with most of the ARGs (Fig. 3).

278

279 **4. DISCUSSION**

280 Some antibiotic resistance features in urban wastewater, such as the most common genes and
281 predominant bacterial phyla seem to follow a general pattern (Gatica et al. 2016; Manaia et al.
282 2016; Narciso-da-Rocha et al. 2018). However, it is recognized that each UWTP has

283 specificities due not only to biogeographic factors but also to the composition of the influents
284 received, the age of the plant, among others. These factors, may surpass the effect of climate
285 conditions and lead to final effluents with distinct resistance features and community
286 composition. The bacterial community composition may be an important driver in determining
287 the prevalence and patterns of resistance in the final effluent. Indeed, bacteria selection is
288 suggested as an important biological process ruling the fate of ARGs during wastewater
289 treatment, eventually with higher impact on the loads of resistance than horizontal gene transfer
290 (Vaz-Moreira et al. 2014; Bengtsson-Palme et al. 2016; Manaia et al. 2016). In this study, we
291 were interested in comparing the bacterial community and resistance profile in final effluents of
292 three UWTPs. Based on such a comparison we aimed at inferring about possible associations
293 between antibiotic resistance and bacterial community members. Assuming that geographical
294 and socio-economic factors may have a strong effect on the bacterial communities and
295 antibiotic resistance loads in the wastewater effluents, three UWTPs located within a distance of
296 7 km were selected for this comparative study. Interesting differences in these plants are the
297 year of construction (more than 25 years ago for PT1, 15 years for PT2, and 18 years for PT3),
298 the reception of hospital effluents (known only for PT2), and the existence of UV disinfection
299 available at the time of sampling only in PT3. In all UWTPs effluents the predominant bacterial
300 phyla were the same, confirming previous reports that highlight that wastewater samples hold
301 similar bacterial community compositions, at high taxonomic ranks (Munck et al. 2015). As in
302 previous reports, *Proteobacteria* and *Bacteroidetes* were among the major phyla, although in
303 Portugal wastewater samples, *Actinobacteria* are consistently poorly represented (Ye and
304 Zhang 2013; Munck et al. 2015; Binh et al. 2018; Narciso-da-Rocha et al. 2018) (Fig. 1).
305 However, eventually as a result of the history and functioning differences of the three UWTPs,
306 the bacterial community composition clustered into three distinct groups, each represented by a
307 UWTP. Therefore, the current study evidenced that UWTPs located in close regions discharge
308 treated effluents with a distinct profile of taxonomic groups (Fig. 1 and Fig. 2). Indeed, it was
309 observed that in a total of 9 to 12 samples collected in each UWTP, in distinct seasons, with
310 average temperatures 9 °C (early Spring) and 17 °C (early Autumn), the final effluent of each
311 plant fell into the same group, with the effluents of the three UWTPs forming three distinct
312 groups. The UWTPs effect was more notorious than the season effect, for which no significant

313 differences were observed. This suggests that the microbiota present in the final effluent is
314 somehow characteristic of a given UWTP. This may be related to the previously noted
315 functional stability of wastewater treatment bioreactors, which permit the maintenance of the
316 system, buffering the occurrence of possible perturbations (LaPara et al. 2002).

317 UWTP PT1 was, among the three analysed, the one with the most distinct bacterial community,
318 with a significantly higher relative abundance of members of the families [*Weeksellaceae*],
319 *Flavobacteriaceae*, *GZKB119*, *Porphyromonadaceae*, *Comamonadaceae*, and
320 *Pseudomonadaceae* (Fig. 2A, Fig. 3). Also, of note was the fact that the phyla *TM7* (Candidatus
321 *Saccharibacteria*) and *OD1* (Candidatus *Parcubacteria*) were more abundant in PT2 and PT3
322 than in PT1, eventually as a result of the higher bacterial diversity. Members of these phyla
323 have small genomes and reduced metabolic capabilities (Albertsen et al. 2013; Nelson and
324 Stegen 2015), although yielding several genes involved in complex carbon degradation or
325 sulfate reduction (Kantor et al. 2013), which may explain their higher prevalence in the most
326 recent UWTPs (Zhang et al. 2012; Ye and Zhang 2013).

327 Except for *bla*_{OXA-A}, with the highest abundance and prevalence values in PT1, the average
328 values of abundance or of prevalence of ARGs and *intl1* did not differ by more than 1.2 log-unit
329 in the different final effluents. This observation suggests that the average loads of ARGs may
330 not differ much in different UWTPs located in the same region. However, the load of ARGs and
331 *intl1* suffered variations over the distinct sampling campaigns, and these were observed to be
332 wider than those registered for the bacterial community composition (Fig. 2A vs. Fig. 4B).

333 Interestingly, the distribution of the relative abundance of bacterial community members and of
334 the ARGs and *intl1* gene prevalence coincided, with the different UWTPs originating distinct
335 groups (Fig. 2A vs. Fig. 4B). This observation seems to confirm the hypothesis that the selection
336 promoted by wastewater treatment, rather than only horizontal gene transfer, may contribute to
337 explain the occurrence of ARGs in the final effluent.

338 Bacterial families whose prevalence might be correlated with that of ARGs and *intl1* included
339 *Aeromonadaceae* and *Campylobacteraceae* (*Proteobacteria*), and *Veillonellaceae* (*Firmicutes*),
340 whose relative abundance was significantly positively correlated with ARGs prevalence (Fig. 3).
341 Surprisingly, none of these bacterial families was correlated with the *sul1* gene, one of the most
342 widespread genes in the environment. Except for *Aeromonadaceae*, these bacterial groups are

343 not among the most probable harbours of the analysed ARGs. However, these bacterial
344 lineages may represent bacterial groups that co-exist with the ARGs hosts, in particular
345 because they share similar physiological properties or ecology traits. However, it must be
346 argued that probably groups with relative abundance <1%, whose tracking is difficult with the
347 technique used to characterize the bacterial community, may, eventually, represent the most
348 important carriers of ARGs in wastewater habitats. Further studies based on epicPCR or long-
349 read sequence analyses may bring new insights into this issue (Spencer et al. 2016; Manaia et
350 al. 2018).

351

352 **5. CONCLUSIONS**

353 The microbiota present in the final effluent is somehow characteristic of a given UWTP and it is
354 not strongly influenced by sampling date or season. In a UWTP, ARGs prevalence presents
355 higher variation than the relative abundance of bacterial families. The average ARGs loads may
356 not differ sharply in different UWTPs located in the same region. And, groups such as
357 *Aeromonadaceae*, *Campylobacteraceae*, *Veillonellaceae*, *Weeksellaceae*, and
358 *Porphyromonadaceae* were observed to be significantly positively correlated with some ARGs.

359

360 **ACKNOWLEDGMENTS:**

361 The authors gratefully acknowledge the support of the staff of the UWTPs and supplying entities
362 that made this study possible by providing the water samples; Christophe Merlin that kindly
363 provided the integrase gene qPCR protocol and Gonçalo Macedo and Jaqueline Rocha for
364 technical assistance on sampling, DNA extraction, and qPCR analysis.

365 This work was funded by National Funds from FCT – Fundação para a Ciência e a Tecnologia
366 through project WaterJPI/0001/2013 STARE – “Stopping Antibiotic Resistance Evolution” and
367 UID/Multi/50016/2013. IVM was supported by the FCT grant (SFRH/BPD/87360/2012).

368

369 **Conflict of Interest statement:** The authors declare that they have no conflict of interest.

370

371 **REFERENCES:**

- 372 Albertsen M, Hugenholtz P, Skarshewski A, et al (2013) Genome sequences of rare, uncultured
373 bacteria obtained by differential coverage binning of multiple metagenomes. *Nat*
374 *Biotechnol* 31:533–538. doi: 10.1038/nbt.2579
- 375 Becerra-Castro C, Lopes AR, Teixeira S, et al (2017) Characterization of bacterial communities
376 from Masseiras, a unique Portuguese greenhouse agricultural system. *Antonie van*
377 *Leeuwenhoek, Int J Gen Mol Microbiol* 110:665–676. doi: 10.1007/s10482-017-0833-7
- 378 Bengtsson-Palme J, Hammarén R, Pal C, et al (2016) Elucidating selection processes for
379 antibiotic resistance in sewage treatment plants using metagenomics. *Sci Total Environ*
380 572:697–712. doi: 10.1016/j.scitotenv.2016.06.228
- 381 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful
382 approach to multiple testing. *J R Stat Soc Ser B* 57:289–300. doi: 10.2307/2346101
- 383 Berendonk TU, Manaia CM, Merlin C, et al (2015) Tackling antibiotic resistance: the
384 environmental framework. *Nat Rev Microbiol* 13:310–317. doi: 10.1038/nrmicro3439
- 385 Bernardet JF, Nakagawa Y, Holmes B, et al (2002) Proposed minimal standards for describing
386 new taxa of the family Flavobacteriaceae and emended description of the family. *Int J Syst*
387 *Evol Microbiol* 52:1049–1070. doi: 10.1099/ijs.0.02136-0
- 388 Binh CTT, Petrovich ML, Chaudhary A, et al (2018) Metagenomics Reveals the Impact of
389 Wastewater Treatment Plants on the Dispersal of Microorganisms and Genes in Aquatic
390 Sediments. 84:1–15. doi: <https://doi.org/10.1128/AEM.02168-17>
- 391 Caporaso JG, Kuczynski J, Stombaugh J, et al (2010) QIIME allows analysis of high-throughput
392 community sequencing data. *Nat Methods* 7:335–6. doi: 10.1038/nmeth.f.303
- 393 Chao A (1984) Non-parametric estimation of the number of classes in a population. 11:265–270
- 394 DeSantis TZ, Hugenholtz P, Larsen N, et al (2006) Greengenes, a chimera-checked 16S rRNA
395 gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–
396 5072. doi: 10.1128/AEM.03006-05
- 397 Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*

398 26:2460–2461

399 Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biol Conserv* 1–10. doi:
400 10.1890/0012-9658(2006)87[1465:ATTFHF]2.0.CO;2

401 Figueira V, Serra E, Manaia CM (2011a) Differential patterns of antimicrobial resistance in
402 population subsets of *Escherichia coli* isolated from waste- and surface waters. *Sci Total*
403 *Environ* 409:1017–1023. doi: 10.1016/j.scitotenv.2010.12.011

404 Figueira V, Vaz-Moreira I, Silva M, Manaia CM (2011b) Diversity and antibiotic resistance of
405 *Aeromonas* spp. in drinking and waste water treatment plants. *Water Res* 45:5599–5611.
406 doi: 10.1016/j.watres.2011.08.021

407 Gatica J, Tripathi V, Green S, et al (2016) High Throughput Analysis of Integron Gene
408 Cassettes in Wastewater Environments. *Environ Sci Technol* 50:11825–11836. doi:
409 10.1021/acs.est.6b03188

410 Ju F, Li B, Ma L, et al (2016) Antibiotic resistance genes and human bacterial pathogens: Co-
411 occurrence, removal, and enrichment in municipal sewage sludge digesters. *Water Res*
412 91:1–10. doi: 10.1016/j.watres.2015.11.071

413 Kantor RS, Wrighton KC, Handley KM, et al (2013) Small genomes and sparse metabolisms of
414 sediment-associated bacteria from four candidate phyla. *MBio* 4:. doi:
415 10.1128/mBio.00708-13

416 LaPara TM, Nakatsu CH, Pantea LM, Alleman JE (2002) Stability of the bacterial communities
417 supported by a seven-stage biological process treating pharmaceutical wastewater as
418 revealed by PCR-DGGE. *Water Res* 36:638–646. doi: 10.1016/S0043-1354(01)00277-9

419 Leps J, Smilauer P (2014) *Multivariate Analysis of Ecological Data using CANOCO* 5

420 Lozupone C, Knight R (2005) UniFrac : a New Phylogenetic Method for Comparing Microbial
421 Communities UniFrac : a New Phylogenetic Method for Comparing Microbial Communities.
422 *Appl Environ Microbiol* 71:8228–8235. doi: 10.1128/AEM.71.12.8228

423 Manaia CM, Macedo G, Fatta-Kassinos D, Nunes OC (2016) Antibiotic resistance in urban
424 aquatic environments: can it be controlled? *Appl Microbiol Biotechnol* 100:1543–1557. doi:

425 10.1007/s00253-015-7202-0

426 Manaia CM, Rocha J, Scaccia N, et al (2018) Antibiotic resistance in wastewater treatment
427 plants: Tackling the black box. *Environ Int* 115:312–324

428 Michael I, Rizzo L, McArdell CS, et al (2013) Urban wastewater treatment plants as hotspots for
429 the release of antibiotics in the environment: a review. *Water Res* 47:957–995. doi:
430 10.1016/j.watres.2012.11.027

431 Moreira NFF, Narciso-da-Rocha C, Polo-López MI, et al (2018) Solar treatment (H₂O₂, TiO₂-
432 P25 and GO-TiO₂ photocatalysis, photo-Fenton) of organic micropollutants, human
433 pathogen indicators, antibiotic resistant bacteria and related genes in urban wastewater.
434 *Water Res*

435 Munck C, Albertsen M, Telke A, et al (2015) Limited dissemination of the wastewater treatment
436 plant core resistome. *Nat Commun* 6:8452. doi: 10.1038/ncomms9452

437 Narciso-da-Rocha C, Rocha J, Vaz-Moreira I, et al (2018) Bacterial lineages putatively
438 associated with the dissemination of antibiotic resistance genes in a full-scale urban
439 wastewater treatment plant. *Environ Int* 118:. doi: 10.1016/j.envint.2018.05.040

440 Nelson WC, Stegen JC (2015) The reduced genomes of Parcubacteria (OD1) contain
441 signatures of a symbiotic lifestyle. *Front Microbiol* 6:. doi: 10.3389/fmicb.2015.00713

442 Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic
443 and functional profiles. *Bioinformatics* 30:3123–3124

444 Paulson JN, Stine CO, Bravo HC, Pop M (2014) Robust methods for differential abundance
445 analysis in marker gene surveys. *Nat Methods* 10:1200–1202. doi:
446 10.1038/nmeth.2658.Robust

447 Rizzo L, Manaia C, Merlin C, et al (2013) Urban wastewater treatment plants as hotspots for
448 antibiotic resistant bacteria and genes spread into the environment: a review. *Sci Total
449 Env* 447:345–360. doi: 10.1016/j.scitotenv.2013.01.032S0048-9697(13)00042-9 [pii]

450 Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets.
451 *Bioinformatics* 27:863–864. doi: 10.1093/bioinformatics/btr026

452 Shannon CE, Weaver W (1963) The mathematical theory of communication. University of
453 Illinois Press, Urbana

454 Simpson EH (1949) Measurement of Diversity. *Nature*

455 Spencer SJ, Tamminen M V., Preheim SP, et al (2016) Massively parallel sequencing of single
456 cells by epicPCR links functional genes with phylogenetic markers. *ISME J* 10:427–436.
457 doi: 10.1038/ismej.2015.124

458 Vandamme P, De Ley J (1991) Proposal for a New Family, Campylobacteraceae. *Int J Syst*
459 *Bacteriol* 41:451–455. doi: 10.1099/00207713-41-3-451

460 Vaz-Moreira I, Nunes OC, Manaia CM (2014) Bacterial diversity and antibiotic resistance in
461 water habitats: Searching the links with the human microbiome. *FEMS Microbiol Rev*
462 38:761–778

463 Vikesland PJ, Pruden A, Alvarez PJJ, et al (2017) Toward a Comprehensive Strategy to
464 Mitigate Dissemination of Environmental Sources of Antibiotic Resistance. *Environ Sci*
465 *Technol* 51:13061–13069. doi: 10.1021/acs.est.7b03623

466 Wollenberg AL van den (1977) Redundancy analysis an alternative for canonical correlation
467 analysis. *Psychometrika* 42:207–219

468 Wu D, Dolfing J, Xie B (2018) Bacterial perspectives on the dissemination of antibiotic
469 resistance genes in domestic wastewater bio-treatment systems: beneficiary to victim.
470 *Appl Microbiol Biotechnol* 102:597–604. doi: 10.1007/s00253-017-8665-y

471 Ye L, Zhang T (2013) Bacterial communities in different sections of a municipal wastewater
472 treatment plant revealed by 16S rDNA 454 pyrosequencing. *Appl Microbiol Biotechnol*
473 97:2681–2690. doi: 10.1007/s00253-012-4082-4

474 Zhang T, Shao M-F, Ye L (2012) 454 Pyrosequencing reveals bacterial diversity of activated
475 sludge from 14 sewage treatment plants. *ISME J* 6:1137–47. doi: 10.1038/ismej.2011.188

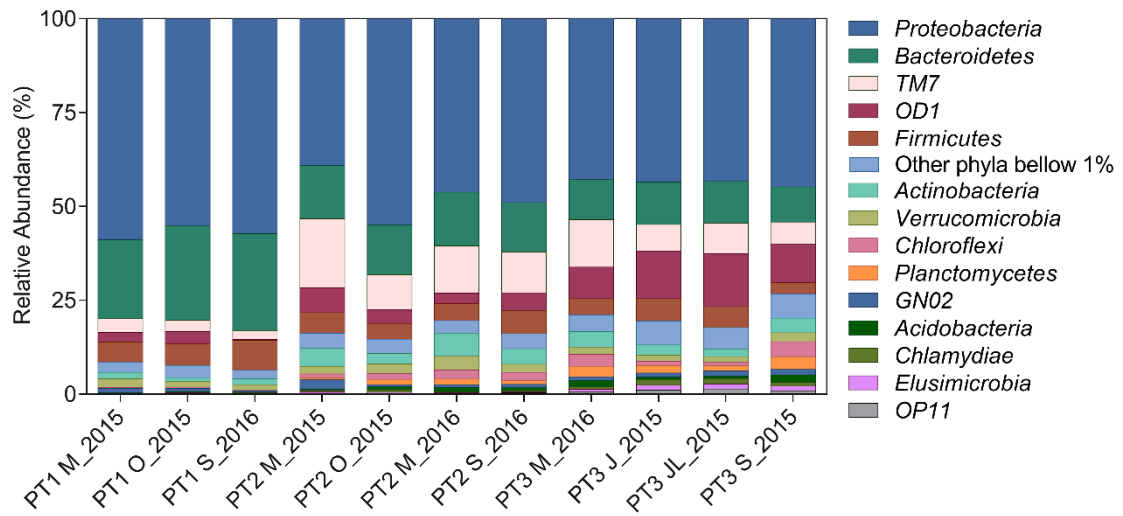
476

477 **Table 1** Range of richness and alpha diversity indices and estimated values (mean value \pm SD)
 478 for the final effluent of different urban wastewater treatment plants (UWTPs), PT1, PT2 and PT3
 479 analysed over the different sampling campaigns.

UWTPs	Observed OTUs	Richness Estimator	Diversity indices		
		Chao1	Shannon	Simpson	PD whole tree
PT1	5796-7996	9024-12053	7.84-8.85	0.969-0.987	201-259
	7238\pm861^a	10994\pm1171^a	8.54\pm0.30^a	0.979\pm0.008^a	236\pm22^a
PT2	4382-8229	7809-11856	8.57-9.42	0.985-0.991	189-294
	6732\pm1114^{a,b}	10799\pm1139^a	8.99\pm0.27^{a,b}	0.989\pm0.002^b	260\pm30^a
PT3	4454-7350	7018-11006	8.74-9.95	0.986-0.996	225-310
	5704\pm996^b	9158\pm1336^b	9.24\pm0.48^b	0.991\pm0.004^b	262\pm27^a

480 Note: Sample PT3.2 S_2015 was excluded from the analyses due to the low number of reads
 481 (n= 37 932). a, b, and c indicate significant differences ($p < 0.01$) between UWTPs.
 482

483 **Figures:**

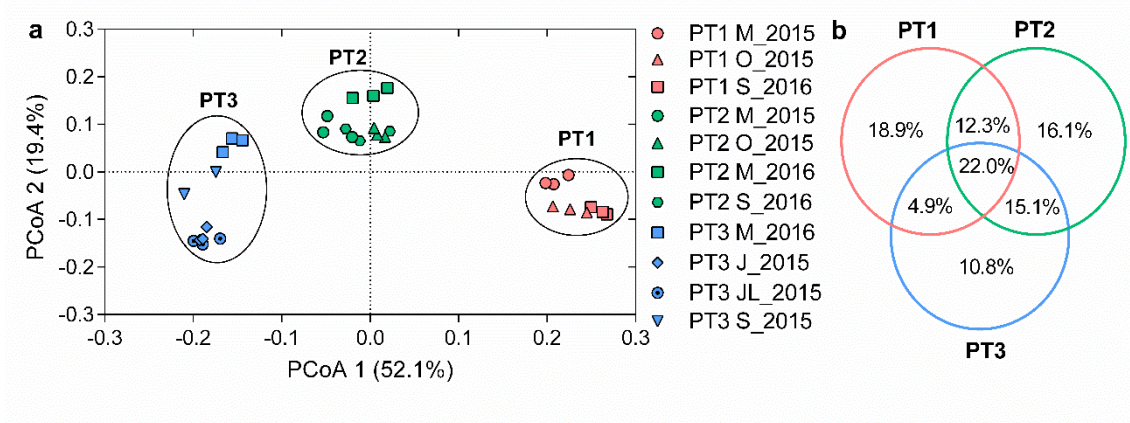


484

485

486 **Fig. 1** Relative abundance at bacterial phyla observed in the final effluent of the urban
487 wastewater treatment plants (PT1, PT2 and PT3), in distinct sampling campaigns, each
488 including three consecutive days, in March 2015 (M_2015), October 2015 (O_2015),
489 March 2016 (M_2016), September 2016 (S_2016), June 2015 (J_2015), July 2015 (JL_2015),
490 and September 2015 (S_2015). Taxa with abundance below 1% in all samples were designated
491 as other phyla.

492



493

494 **Fig. 2** Analysis of the wastewater bacterial diversity. **a)** Principal coordinates analysis (PCoA)

495 based on weighted UniFrac distances between different final effluents of the UWTPs. PT1

496 (pink), PT2 (green), and PT3 (blue), and the respective sampling campaigns: March 2015

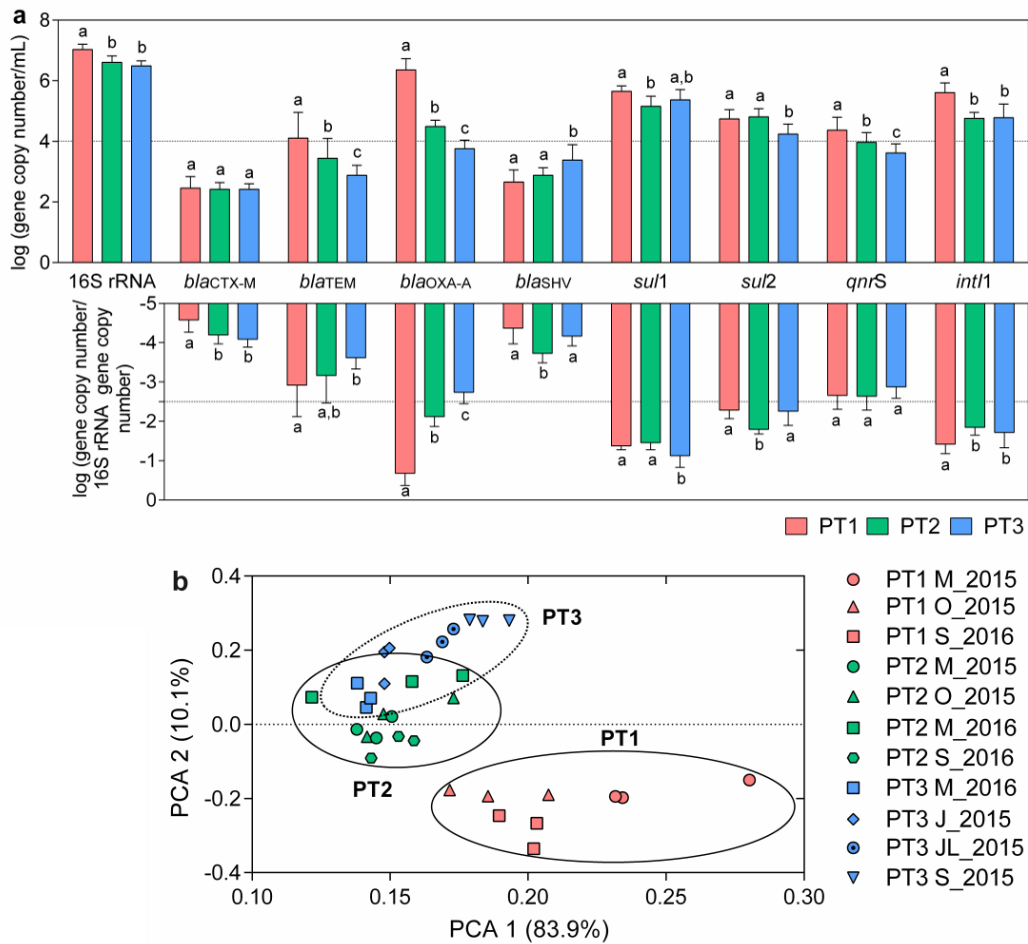
497 (M_2015), October 2015 (O_2015), March 2016 (M_2016), September 2016 (S_2016),

498 June 2015 (J_2015), July 2015 (JL_2015), and September 2015 (S_2015). **b)** Venn diagram

499 representation of the percentage of shared and exclusive bacterial OTUs between the different

500 UWTPs.

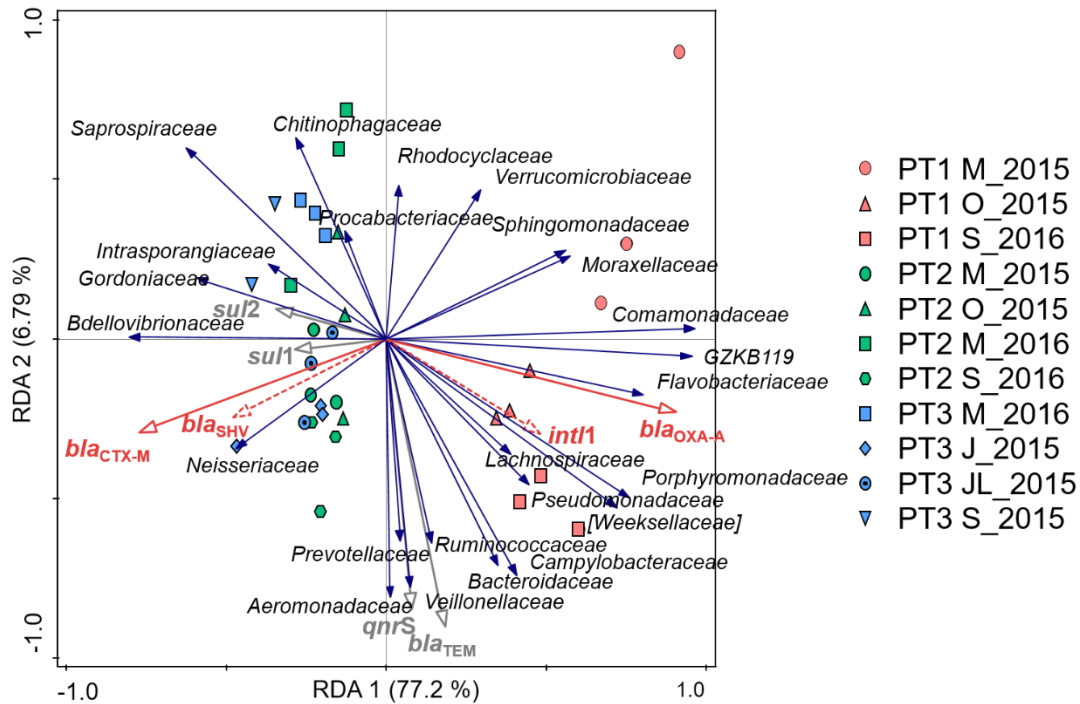
501



511

512 **Fig. 4** Quantification of ARGs and *int1* in the final effluent of the urban wastewater treatment
 513 plants (PT1, PT2, and PT3). **a**) Gene abundance (gene copy number / mL of sample) in the
 514 upper part of the figure or prevalence (target gene copy number / 16S rRNA gene copy number)
 515 in the bottom. a, b and c indicate significantly ($p < 0.01$) different Tukey's groups; **b**) Principal
 516 component analysis of the distribution of ARGs and *int1* gene prevalence.

517



518

519 **Fig. 5** Redundancy Analysis (RDA) triplot of the bacterial community composition at the family
 520 level (relative abundance > 1 %) and environmental variables (ARGs and *int1* prevalence) in
 521 the 32 final effluent samples. Blue arrows indicate the members of the bacterial community. The
 522 grey arrows represent environmental variables (ARGs and *int1*) with no significant correlation,
 523 based on the Monte Carlo permutation test after FDR correction. The pointed red and red
 524 arrows show the variables with significance lower than 0.05 and 0.01, respectively. PT1 (pink),
 525 PT2 (green), and PT3 (blue) regards the UWTPs in distinct sampling campaigns: March 2015
 526 (M_2015), October 2015 (O_2015), March 2016 (M_2016), September 2016 (S_2016),
 527 June 2015 (J_2015), July 2015 (JL_2015), and September 2015 (S_2015).