Universitat Autònoma de Barcelona

Insights into the microbial community of an enhanced biological phosphorus removal plant

JÚLIA GRANDE MARTÍ

MÀSTER EN ENGINYERIA BIOLÒGICA I AMBIENTAL

JUAN BAEZA, ALBERT GUISASOLA & ELOI PARLADÉ

07/02/2025

ABSTRACT

The biological phosphorus removal (EBPR) process relies mainly on polyphosphate accumulating organisms (PAOs). Nowadays their characterization still remains a challenge due to the complexity of wastewater communities and the lack of a single methodology that accurately identifies and quantifies these organisms. Fluorescence in situ hybridization enables identification/quantification of specific targeted groups, whereas next sequencing generation techniques provide a broader taxonomic overview of the microbial community. However, across literature there are significant discrepancies between these methodologies results, highlighting the need of a systematic approach reliable for PAO characterization.

In this study three bacteria genera exhibiting PAO phenotype are evaluated. The three show differences in their relative abundances comparison, being *Ca*. Accumulibacter the most divergent (abundances <1% in FISH vs. >28% in NGS).

Keywords: Enhanced biological phosphorus removal (EBPR), Polyphosphate-accumulating organisms (PAO) Ca. Accumulibacter, FISH, diversity.

1. INTRODUCTION

Phosphorus (P) is a vital nutrient for all living organisms, but its cycle has been largely altered by human activities, resulting in increasing concerns about its supplies (Liu et al., 2016). It is a limited and non-renewable source, and it is estimated that the main source of P could be depleted in the next 50-100 years (Cordell et al., 2009) .

Moreover, since it is a key nutrient that stimulates the growth of algae and other photosynthetic microorganisms, it must be removed from wastewater to avoid eutrophication in aquatic water systems (Oehmen et al., 2007). To prevent eutrophication, regulatory limits have been established for the quantity of P present in the effluent of wastewater treatment plants (WWTPs) discharged into water bodies (Ruiz-Haddad et al., 2024). For example, the European Environment Agency requires an 80% reduction in P release or a limit of 1-2 mg L^{-1} in the effluent of WWTPs (European Environment Agency, 2019). The European directive for wastewater treatment is currently being reviewed, the limits for P concentration in the effluent will be more stringent for certain specific areas (European Parliament, 2024).

This scenario explains why the effort of recovering P from wastewater is increasing exponentially. Wastewater contains a lot of phosphorus, which can be recovered. Only accounting human excreta, it is estimated to contain 3-4 million metric tons of P per year (Mihelcic et al., 2011; Nagy et al., 2019). Among that, urine contains a high concentration of P (Illakwahhi et al., 2024); it was reported that recovering P from domestic wastewater could meet 15-20% of the global demand for P (Nagy et al., 2019).

Nowadays P removal from wastewater is often achieved by implementing the enhanced biological phosphorus removal (EBPR) process. It is the most cost-effective, efficient and sustainable method for P removal from wastewater (Kulakovskaya et al., 2012; Pepin & Wood, 1986; Seviour et al., 2003). EBPR promotes the removal of P from wastewater without the need for chemical precipitants. This process modifies the activated sludge process by recirculating sludge through anaerobic and aerobic conditions (Barnard, 1975). The P-rich biomass can be used directly as a fertilizer or the P can be released as orthophosphate and precipitated in form of struvite or vivianite (Yuan et al., 2012).

The success of EBPR system relies on the activity of a specialized group of microorganisms known as polyphosphate-accumulating organisms (PAO). These bacteria have the ability to take up and store P as intracellular polyphosphate granules (Cyprowski et al., 2018). Under anaerobic conditions, PAO can take up carbon sources such as volatile fatty acids (VFAs) and store them as carbon polymers (poly-*β*-hydroxyalkanoates)(Oehmen et al., 2007). The energy to store this polymer is obtained from the hydrolysis of polyphosphate (poly-P) and glycogen. The phosphate concentration in the liquid media increases during the anaerobic phase because of poly-P hydrolysis. During the aerobic (and/or anoxic) phase, the stored PHA is consumed, generating energy for growth, uptake of orthophosphate and replenishment of glycogen

pool (Zhou et al., 2010). Although this is the model metabolism for PAO not all of them exhibit the same behavior. For example, no PHA or glycogen have been identified in situ in *Tetrasphaera (Fernando et al., 2019)*.

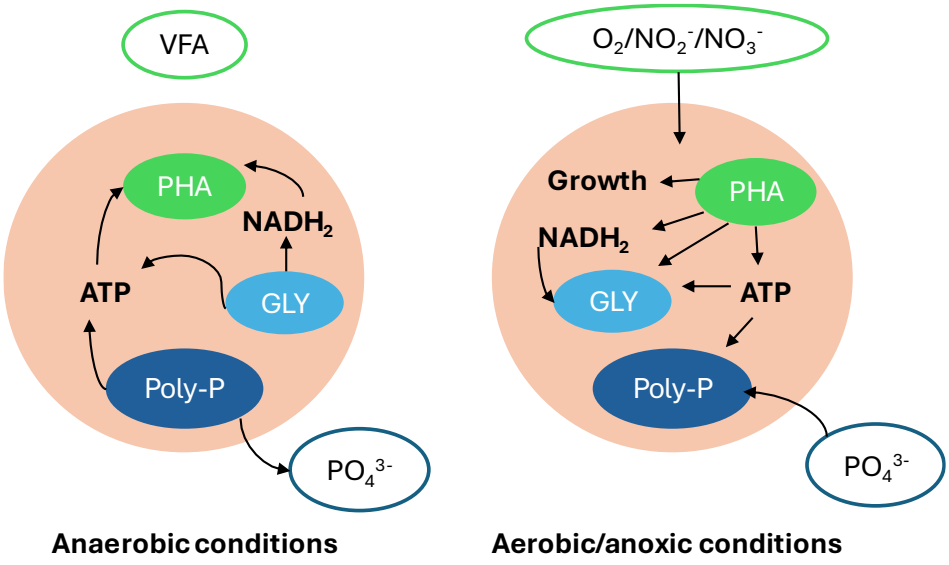

Figure 1. Schematic representation of PAO metabolism.

Activated sludge in the EBPR process comprises thousands of bacterial species, but to date only few genera have been recognized as PAO (Petriglieri, Petersen, et al., 2022). Among the reported PAO (Fig. 1), *Candidatus* Accumulibacter, *Tetrasphaera* and *Candidatus* Dechloromonas were identified as the most significant PAO (Dueholm et al., 2022a). Besides these taxa, other putative PAO are being considered, for example, the genera of filamentous bacteria, *Thiothrix* can store polyphosphate (Wanner et al., 1987) and Rubio-Rincón et al., 2017 stated that it could behave like PAO with a mixotrophic metabolism. There is a range of genera related with PAO activity, but, genome coverage is poor for all them except for *Ca*. Accumulibacter *(Nielsen et al., 2019a)* and their phylogeny has constantly been reviewed. The role and significance of the majority of EBPR related taxa are still to be defined (Rey-Martínez et al., 2019). Figure 2 displays a taxonomic classification of organisms exhibiting PAO phenotype.

Ca. Accumulibacter was the first microorganism identified as PAO, and it is the most studied one. Later on, it was classified into two types I or II (and 9 clades)(Ruiz-Haddad et al., 2024). On the other hand, *Tetrasphaera* genus has been recently reclassified, and some species have been moved to the genus *Phyciococcus*, *Knoellia, Ca.* Phosphoribacter, *Ca*. Lutibacillus (Nouioui et al., 2018; Ruiz-Haddad et al., 2024). Among these, *Ca*. Phosphoribacter has been described as a dominant PAO in EBPR wastewater treatment worldwide (Singleton et al., 2022) (green lines in Fig. 2 represent all the taxa previously categorizes as *Tetrasphaera*). Besides, *Dechloromonas* has been recently renamed *Azonexus*. This genre exhibits a very similar physiology to *Ca*. Accumulibacter, having dynamic levels of poly-P, glycogen and PHA (Petriglieri et al., 2021).

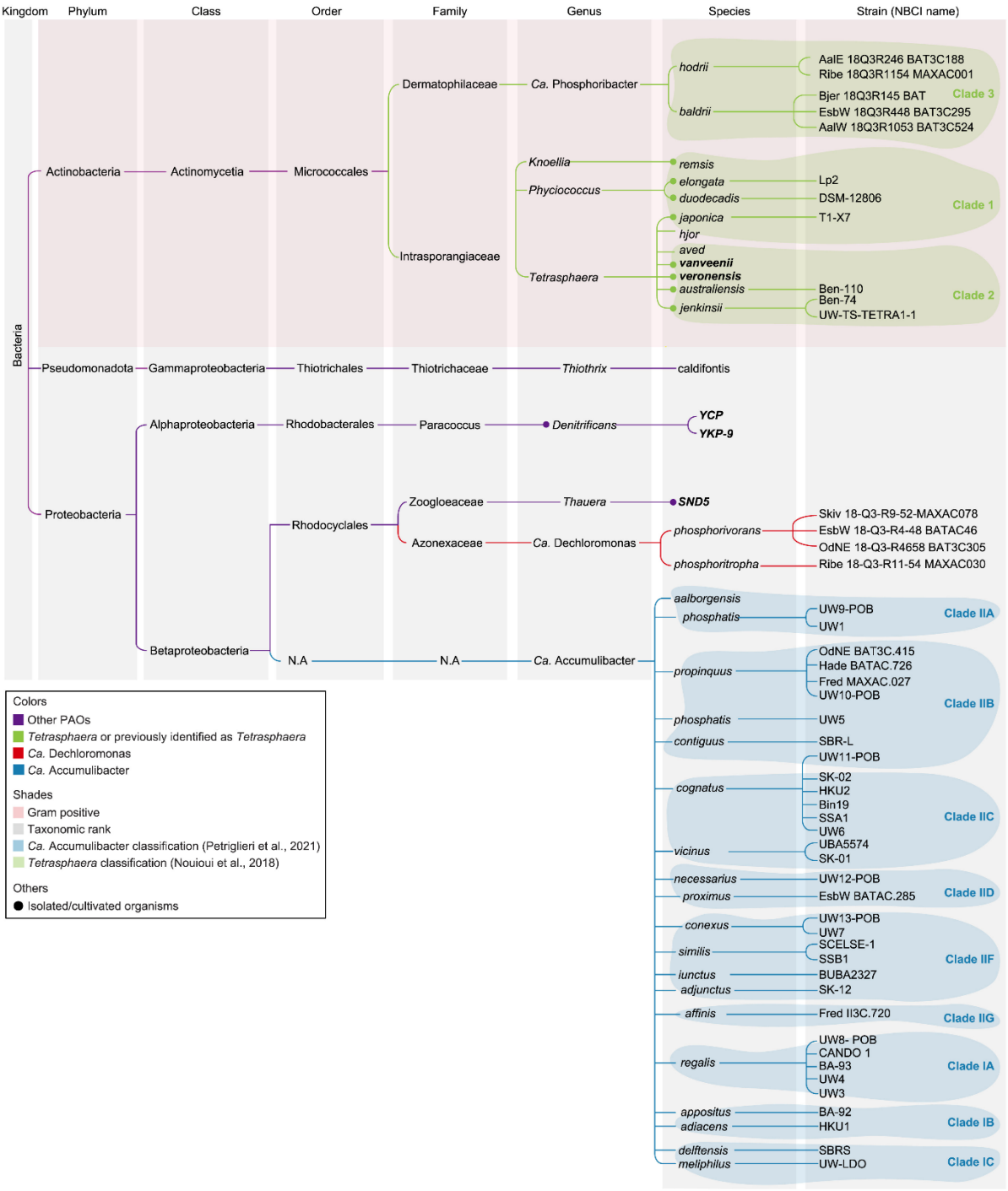

Figure 2. Taxonomic classification of organisms displaying PAO activity. Extracted from (Ruiz-Haddad et al., 2024).

In the EBPR process, glycogen-accumulating organisms (GAO) have been traditionally considered potential competitors to PAOs. These organisms are also able to proliferate under alternating anaerobic and aerobic conditions but, unlike PAO, GAO store carbon as intracellular glycogen without contributing to P removal (Mino et al., 1995; Oehmen et al., 2007). Extensive research has focused on trying to reduce GAO presence in EBPR systems and finding the optimal conditions that favors PAO over GAO (Lopez-Vazquez et al., 2009; Tayà et al., 2013). However, full-scale systems are more complex environments and many of the proposed PAO and GAO organisms are not even present

or only present in low abundances (Stokholm-Bjerregaard et al., 2017a). In full-scale EBPR systems GAO don't appear to be a real problem (Nielsen et al., 2019a).

The characterization of PAO organisms is challenging. Wastewater contains complex and diverse microbial communities, making PAO identification and quantification difficult. Traditionally, microbial wastewater communities have been studied using fluorescence in situ hybridization (FISH), a technique that provides valuable insights into the spatial organization and allows the identification and quantification of the relative abundance of specific microbial groups. However, this technique has limitations when assessing the broader microbial community. In any case, the democratization of metagenomics has opened a new paradigm in the study of these communities. In contrast to FISH, the advantage of next-generation sequencing (NGS), such as 16S rRNA sequencing, is its ability to capture a wider microbial diversity. Unlike FISH, metagenomics offers a culture-independent analysis of the entire microbial community, enabling the identification and characterization of a wild range of microorganisms, both abundant and rare. One key advantage of metagenomics is that it does not require prior knowledge of the microorganisms present, eliminating the need to design experiments around pre-selected probes. The combination of both techniques is commonly employed in microbiological studies. However, there is extensive literature highlighting notable differences in species resolution between the two methods. The techniques can yield differing microbial compositions, not only in relative abundance but also in the presence or absence of certain taxa The reasons for these discrepancies remain unclear, but they are likely due to a combination of factors inherent to both techniques.

The main goal of this work is to study a PAO-enriched population under the novel configuration EBPR + Sludge Side Stream Fermenter using two complementary molecular techniques. To achieve this, a comprehensive bibliographic review will be conducted to understand the differences between the methodologies and to develop hypotheses regarding the potential causes of these differences. The bacterial population under study will be characterized using Illumina sequencing, with a detailed analysis of the bioinformatics steps involved. Simultaneously, the population will be studied through FISH. Finally, a comparative analysis of both outcomes will be performed to identify key differences, and guidelines will be proposed to improve the characterization of PAO-enriched populations for future applications.

2. MATERIALS AND METHODS

2.1 Plant operation

The samples analyzed in this work were extracted from a pilot plant with an A^2/O configuration operating for simultaneous carbon, N and P removal. The research was focused on tracking the organisms able to remove P, but removing N is also essential in waste-water treatment.

The continuous A^2/O configuration is composed of an anaerobic reactor, where P is released and PAO growth is favored, an anoxic reactor, where denitrification occurs and finally the aerobic reactor where nitrifying bacteria oxidize ammonia and P is accumulated by PAO.

A modification of the conventional A^2/O was conducted (S2EBPR) (Fig. 3). A side stream fermenter was added to the external recirculation. That way there is less/no-direct dependence on inlet VFA, it mitigates the effect of excessive nitrate and favors PAO over GAO.

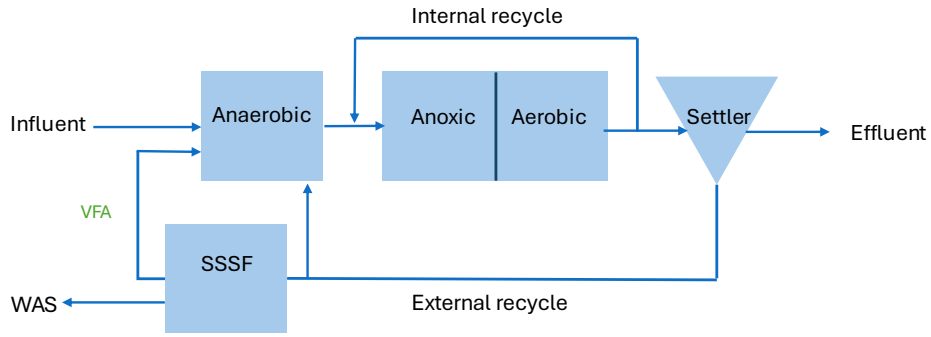

Figure 3. Scheme of the S2EBPR process.

2.2 Sampling

To perform this study 10 activated sludge samples were collected from the pilot plant. The samples were collected during different operation periods and under diverse conditions, as indicated in table 1.

Table 1. Samples parameters of this study obtained from the pilot plant

Sample ID	Wwtype	Reactor	P influent	P reactor (mg/l)	P effluent (mg/L)	Ammonia influent (mg/l)	Nitrate effluent (mg/l)	Nitrite effluent (mg/l)	COD influent (mg/l)	HRT SSF
C800-AER1	real	AER	5.53	0.14	0.16	76.5	18.1	1.1	808	2.5
C800-AER2	real	AER	5.52	0.19	0.16	58.4	3.31	1.46	776	2.5
C800-AER3	real	AER	6.67	0.09	0.17	60	15.09	1.28	791	2.5
C800-SSSF1	real	SSSF	5.53	58.56	0.16	76.5	18.1	1.1	808	2.5
C800-SSSF2	real	SSSF	5.52	62.69	0.16	58.4	3.31	1.46	776	2.5
C800-SSSF3	real	SSSF	6.67	52.14	0.17	60	15.09	1.28	791	2.5
AER-1d	synthetic	AER	11	0.21	0.09	43	7.96	0	650	1
AER-2-5d	synthetic	AER	10	0.32	0.21	40	10.57	0	600	2.5
SSF-1d	synthetic	SSSF	11	67	0.09	43	7.96	0	650	1
SSF-2-5d	synthetic	SSSF	10	210.6	0.21	40	10.57	0	600	2.5

2.3 Fluorescent in situ hybridization (FISH)

FISH technique is a commonly used method to identify PAO in WWTP. It can be used to measure the relative bio abundance of the microbial community. Furthermore, since it maintains the morphology of the cells, it is possible to observe it and have information on the spatial distribution and the equivalent biovolume (Nielsen et al., 2009). It allows single cell level identification.

FISH involves annealing of DNA or RNA probes attached to a fluorescent reporter molecule with specific target sequence of sample DNA, which can be followed under fluorescence microscopy (Shakoori, 2017). The detection target is 16S rRNA, the component of the small subunit of procaryotic ribosomes that targets the sequence Shine-Dalgarmo. The 16S RNAr gene is widespread used in bacteria phylogenetic studies because the sequence is highly conserved among different procaryotic groups (Abellan et al., 2021).

The probes used consist of short DNA sequences (15-25nt) that are designed to hybridize their complementary target sequence on the rRNA structures in the target cell. From the composition of the probe, it is possible to design it to specifically target a narrow phylogenetic group (down to the species level) or any other higher phylogenetic hierarchical. No probes will hybridize to those cells without target sequences. Cells containing the target sequence will on the other hand retain the hybridized probe and due to the large number of ribosomes within active cells thus become fluorescently labeled (Nielsen et al., 2009). Quantitative FISH (qFISH) can be performed, biovolume fractions of individual taxa are calculated as a percentage area of the total biovolume, hybridizing generic bacterial probes (EUBmix probes), that also hybridize with the specific taxa probe.

A typical flowchart for FISH in activated sludge is shown in Fig. 4 extracted from Nielsen et al., 2009. The procedure includes sampling, possible pretreatment, fixation, possible enzyme treatment, hybridization, epifluorescence microscopy, and possible image analysis.

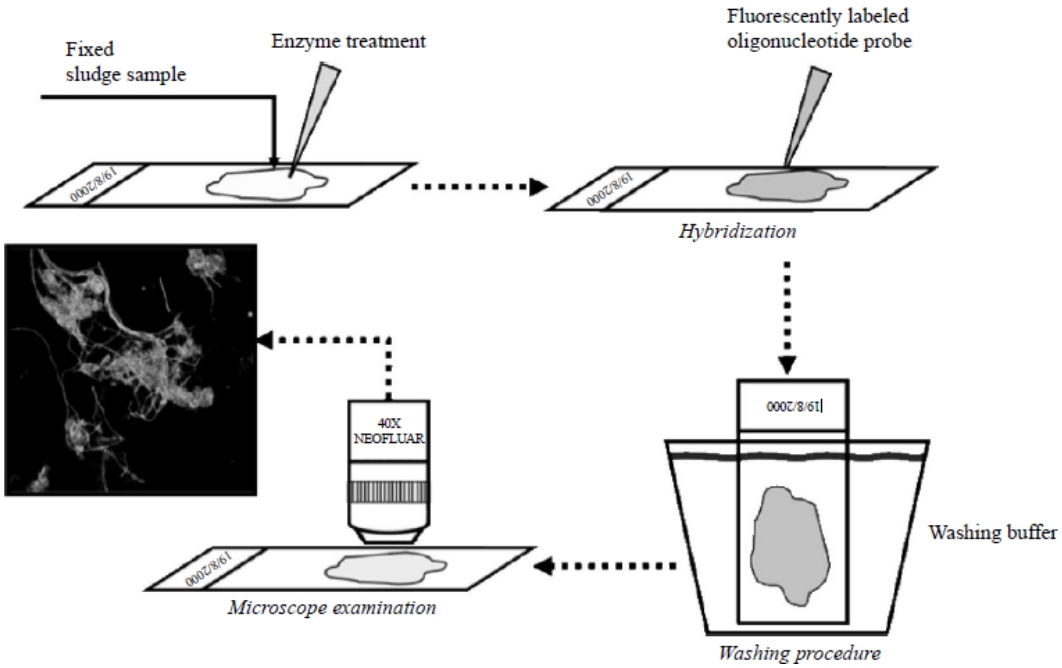

Figure 4. Schematic illustration of the FISH procedure, extracted from Nielsen et al., 2009.

Probe selection

The selection of probes (Table 2) was guided by the preliminary results from the Illumina analysis. Due to time constraints, it was not possible to evaluate all putative PAO. Instead, the most abundant ones (based on preliminary Illumina results) were prioritized, as they were the most likely to produce a detectable signal under the microscope.

Traditionally, three FISH probes (PAOmix, a combination of PAO462, PAO651, and PAO846) have been used to detect PAO. It was considered the most sensitive and specific probe for targeting *Ca*. Accumulibacter (Crocetti et al., 2000; Ruiz-Haddad et al., 2024). However, later research revealed that PAOmix is inadequate to specifically distinguish *Ca*. Accumulibacter from other phylogenetically related taxa. For example, it targets GAO species from the genus *Propionivibrio (Albertsen et al., 2016a)*. This unspecific hybridization can lead to overestimated abundances. It is recommended to employ PAO651 alone to ensure accuracy (Albertsen et al., 2016a).

In 2007, BET135 became the first FISH probe for *Dechloromonas* detection (Kong et al., 2007). BET135 targets Azonexus phosphoritropha and Dech443 targets *Azonexus phosphorivorans* with good specificity and coverage (Petriglieri et al., 2021). The combination of these two probes covers the genera diversity (Mcilroy et al., 2016). However, these FISH probes are not species-specific and may hybridize with other *Dechloromonas* species not associated with PAO activity (Ruiz-Haddad et al., 2024).

As for *Thiothrix,* probe G123T targets a wide range of the isolated taxa; it was designed to specifically hybridize to a signature region on the 16S rRNA of all recognized *Thiothrix* species (Kanagawa et al., 2000).

Table 2. FISH probes used

Probe	**Target group**	**Sequence (5'-3')**	**FA (%)**	**Fluorochrome 5'**
EUB338 (EUBmix)	All bacteria	GCTGCCTCCCGTAGGAGT	60	Alexa fluor 488
EUB338 II (EUBmix)	All bacteria	GCAGCCACCCGTAGGTGT	60	Alexa fluor 488
EUB338 III (EUBmix)	All bacteria	GCTGCCACCCGTAGGTGT	60	Alexa fluor 488
PAO651	*Ca*. Accumulibacter	CCCTCTGCCAAACTCCAG	35	Alexa fluor 647
Bet135	*Ca*. Azonexus phosphoritropha (former Dechloromonas)	ACGTTATCCCCCACTCAATGG	45	Alexa Fluor 594
Dech443	*Ca*. Azonexus phosphorivorans (former Dechloromonas)	ACCCATGCATTTTCTTCCCGG	35	Alexa Fluor 594
G123T	*Thiothrix*	CCT TCC GAT CTC TAT GCA	40	Alexa fluor 555

In order to be able to observe the probes, they need to be labelled with fluorochrome. The selection of fluorochrome is important and several factors must be considered. Firstly, the fluorochromes chosen have to match the excitation wavelengths of the lasers in the microscope. In addition, when using multiple probes, the emission spectra has to be different to avoid spectral overlapping. Furthermore, the extinction coefficient is another important factor, it is a measure of how strongly a fluorochrome absorbs light at a given wavelength. The larger the extinction coefficients are, the stronger signals can be obtained (Daims et al., 2019), which is especially crucial when the fluorescence signals are low. The selected fluorochromes, characteristics and their spectral profiles are shown in table 3 and figure 5, respectively.

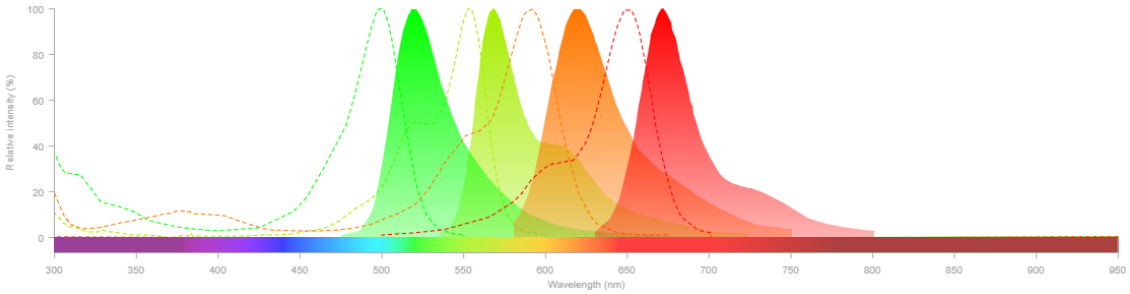

Figure 5. Spectral Profiles of Fluorochromes obtained with Fluorescence SpectraViewer

Table 3. Spectral properties of Alexa Fluor dyes (Termofisher)

Flurochrome	Colour	Max Excitation (nm)	Max Emission (nm)	Extinction coeficient (M–1cm–1)
Alexa Fluor 488	Green	495	519	71,000

Alexa Fluor 555	Yellow-Orange	555	565	155,000
Alexa Fluor 594	Orange-Red	590	617	92,000
Alexa Fluor 647	Far-red	650	668	270,000

FISH protocol

The following protocol was adapted with minimum changes from Daims et al., 2019.

1. Sample collection and fixation

The samples were collected from the pilot plant. Two of the samples from table 1 were evaluated: C800-SSSF1 and SSS-2,5 (distinguished by the type of wastewater) (the fixation was performed for all the samples indicated in table 1).

The microbial biomass was fixed by adding 3 volumes of 4% PFA solution to one volume of sample, and held at 4ºC for 1-3 hours. The samples were centrifuged at 5000 g and the fixative was removed. Afterwards, the cells were washed twice with 1M PBS. For each wash the samples were centrifuged and the supernatant removed. Finally, the cells were resuspended 1 mL of a solution containing 0,5 ml of 1M PBS and 0,5 mL of ice-cold ethanol. The fixed samples can be stored at -20ºC for several months.

2. Sample preparation and immobilization on glass slides

The stored sample were applied directly to the pretreated silanized slides. Silanized slides are used to enhance the cellular adhesion to the slide Surface. The silanization treatment creates covalent bonds between the slide and the cellular structures, ensuring that samples will remain in place throughout the procedure.

Of each sample, 3 preparations were made. First slide with two separated drops of 20 µl (sample type A), second slide with 50 µl of extended sample (sample type B), and third slide with 100 µl of extended sample diluted 1:10 with MiliQ water (sample type C). The side of the pipette tip was used to spread out the sample.

The samples were dried at 37ºC for 20 minutes. The dried biomass was covered with a thin layer of agarose to avoid loss of biomass during subsequent steps. The slides were dipped into warm dissolved agarose (0,5%) and cooled horizontally on ice.

3. Dehydration

Dehydration of the immobilized sample removes water in order to increase the resolution during microscopy. Serial treatments with increasing ethanol concentration efficiently remove the water from the sample. The slides were dipped in a container with 50% ethanol for 3 minutes, followed by 3 min in 80% ethanol and by 3 min in 96% ethanol. The slides were allowed to air dry.

4. Permeabilization by enzymatic treatment

Some cell types require additional treatment by enzymes or chemicals for sufficient permeabilization over the cell wall. Since the nature of our microorganisms was diverse, lysozyme treatment was applied. Other enzymatic or chemical treatments can be applied, e.g. protease K, mutanolsyn, acid, achromopeptidase... But digestion with enzymes, too long incubations or treatment with too harsh chemicals is critical and will cause lysis of some more sensitive cells. So it is often incompatible to access multiple organisms in complex environments. Therefore, the gentlest treatment was chosen.

Lysozime was dissolved to a final concentration of 10mg/mL in 0,05M EDTA, 0,1M Tris-HCl. 10–15 mL of cold lysozyme preparation per slide were applied. The slides were placed in a horizontal position, in a 50 mL Falcon tube, containing tissue paper moistened with dH2O. The slides were incubated for ~ 40 min (10–60 min) at 37℃. Thes slides were washed three times in dH2O, then once in absolute ethanol, and air-dried.

5. Hybridization

2 mL of hybridization buffer for each percentage of formamide were prepared. Formamide concentration modulates the stringency of the hybridization step. Probes requiring different formamide concentrations cannot be applied together. Subsequent hybridizations must be done starting with the highest formamide concentration. Regarding table 1 hybridization buffer was prepared at 40 and 60% of FA.

Table 4. Buffers composition

Buffer type	FA %	FA (µl)	Mili-Q (µl)	5M NaCl (µl)	1 M TrisHcl (µl)	10% SDS (µl)	0.5M EDTA (µl)
Hybridization	40	800	800	360	40	2	-
Hybridization	60	1,200	400	360	40	2	-
Washing	40	-	-	460	1,000	50	500

8 µl of hybridization buffer were transferred onto the slide. 1 µl of each gene probe (work solution 50ng/ µl) were added and carefully mixed with the hybridization buffer. When more than one gene probe is added the quantity remains the same, 1 µl each probe per slide.

The slides were placed horizontally into a 50 mL falcon tube with a piece of tissue moistened with MiliQ water. The tubes were placed in the hybridization oven at 46ºC for 2 hours (minimum time required of 1½ hours).

During hybridization washing buffer was prepared. The same washing buffer was used for the two washing steps (FA 40%). The buffer was prepared in a 50 mL polyethylene tube (x20 times table 3 composition) and preheated at 46ºC.

A few millimeters of the preheated washing buffer were poured over the sample to replace the hybridization buffer (must be carried out in the fume hood to avoid toxic formamide fumes). The slides were transferred into a recipient covered with washing buffer and incubated 15min at 46ºC.

The slides were removed from the washing buffer, dipped in cold MiliQwater and air-dried.

Finally, an antibleaching agent (ProLong Gold Antifade) was added directly to the samples, extended, and covered with a cover glass.

6. Microscope observation and quantification

A Zeiss LSM980 confocal microscope was used for biomass observation. To perform a preliminary assessment of the samples, z-stacks images were taken (different z depths for the same x/y plane). For taxa quantification, 20-30 images were taken for each A preparation by moving the slide in the x/y dimensions and randomly adjusting the focal plane of the microscope in the z dimension. The magnification used was low, 63x with a 1.5x zoom, to capture enough biomass per image and improve statistical accuracy.

Image analysis was performed with *daime (Daims et al., 2006),* a program developed specially for analyzing microbial cell images obtained with fluorescence labeling techniques. The images are segmented to distinguish cells from background prior to the biovolume fraction measurement. Biovolume fraction is calculated as follows, being A^P the area of probe-target population and A^B the area of all bacteria, measured in n pairs.

$$\frac{\sum_{i=1}^{n} A_i^P}{\sum_{i=1}^{n} A_i^B}$$

2.4 Next generation sequencing

NGS involves parallel processing of millions of fragments DNA/RNA simultaneously generating an enormous volume of data in a relatively economic way (Hu et al., 2021).

There are different sequencing platform technologies, however, Illumina is one of the more used NGS methods. The Illumina sequencing platform works as follows (Fig. 6) (adapted from Young & Gillung, 2020). The first step is library preparation, DNA is denaturized into 200-700 bp fragments and the adapters are ligated to the DNA fragments. Then the key step is cluster amplification:

3 The DNA fragments are added to a flow cell and the adapters bind to their complementary oligos, that way DNA fragments attach.
4 The reverse strand of the DNA is synthesized, the double strand is denaturized, and the original strand is eliminated, leaving the complementary stand attached to the flow cell.
5 The strands are amplified through bridge amplification. The ssDNA folds over forming a bridge and the complementary strand is synthesized. The bridge is denaturized leaving two ssDNA bind to the platform.
6 The cycle is repeated many times simultaneously for millions of clusters. The result is clonal amplification of the original fragments. The added nucleotides are labelled with a chemically cleavable fluorescent group at the 3'-OH end, allowing only single

base incorporation in each cycle. After each round of amplification lasers are passed over the flow cell to activate the fluorescence (which is detected and recorded). This process generates a sufficiently strong signal for detection during sequencing.

7 Reverse fragments are eliminated leaving only forward strands to ensure that all the reads come from the same direction.

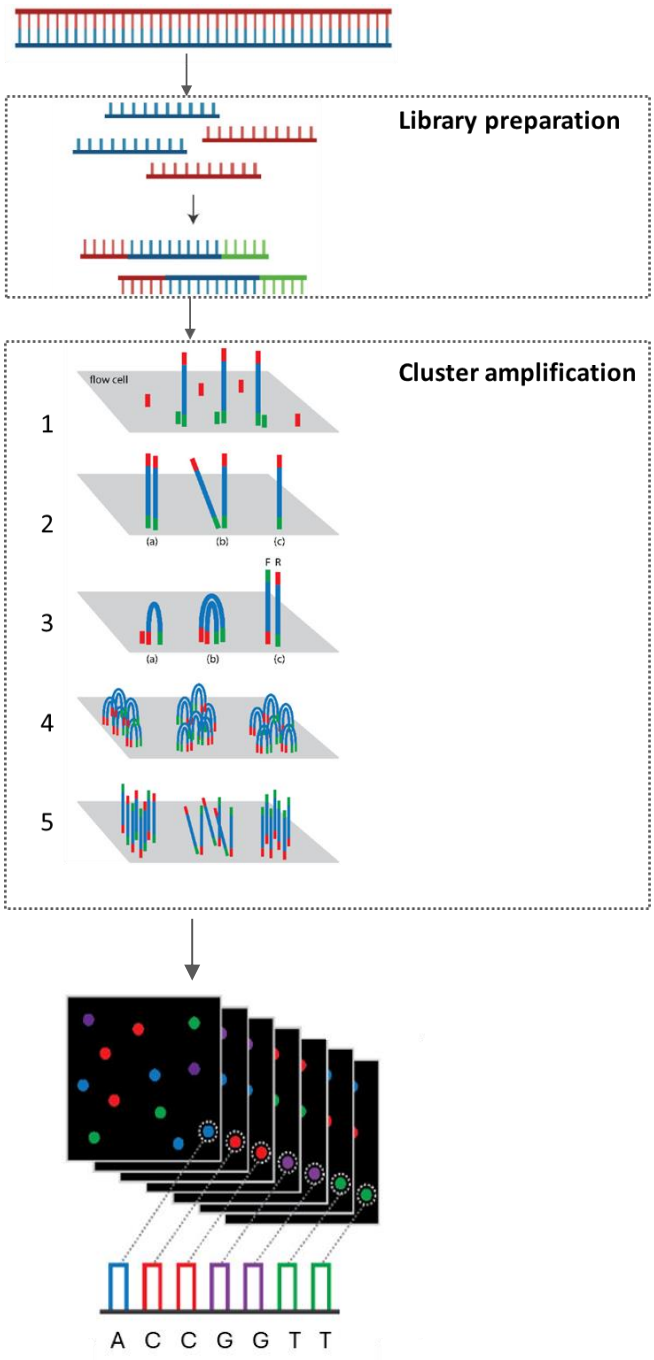

Figure 6. Illumina next- generation sequencing, modified from (Young & Gillung, 2020).

The DNA was extracted following the DNA Isolation Plus kit protocol (Norgen Biotek). The extracted DNA was quantified by a DNA NanoDrop 1000 Spectrophotometer (Waltham, MA, USA), and analyzed by the "Genomic and Bioinformatics service center" at Autonomous university of Barcelona (Barcelona, Spain).

The DNA was amplified using the V3-V4 16S region with the following primer sequences: forward 5'CCTACGGGNGGCWGCAG3' and Reverse 5'GACTACHVGGGTATCTAATCC3'. The amplified product of the first PCR was purified using CleanNGS magnetic beads, which remove excess primers, nucleotides, and other impurities. Samples were indexed in a second PCR round using the IDT for Illumina DNA/RNA UD Indexes and purified using magnetic beads. Indexing allows multiplexing of samples during sequencing by tagging each library with a unique identifier. The library amplicon size was checked before and after indexing using the Agilent 2100 Bioanalyzer. This step ensures that the library meets size requirements for sequencing. The Illumina MiSeq platform was used to sequence the V3-V4 16S libraries with the MiSeq reagent kit v2 at 2x250 bp (both ends of DNA fragments are sequenced). The 16S Metagenomics App (Basespace, Illumina) was used to perform taxonomic classification of 16S rRNA targeted amplicon reads using the RefSeq RDP 16S v3 database. This database allows accurate identification and classification of bacterial and archaeal taxa. The use of this database is the default for the SGIEB service; however, a subsequent analysis was performed, as indicated in the following section.

Bioinformatics reanalysis

The raw sequencing data was requested to SGiEB to perform a bioinformatic reanalysis. The raw sequencing data obtained from Illumina consisted of paired-end reads: forward (R1) and reverse (R2). These reads undergo a series of preprocessing and analytical steps to ensure the reliability of downstream analyses and to characterize microbial communities effectively.

The first step involved assessing the quality profiles of the forward and reverse reads. Quality profiles provide a visual representation of sequencing accuracy across the length of the reads, highlighting regions of lower confidence. This informs the trimming and filtering criteria applied in subsequent quality control. During quality control, sequences were filtered to remove artifacts and trimmed to retain only high-quality regions. The forward and reverse reads were shortened to 240 and 160 nucleotides, respectively. Reads with undetermined bases or excessive errors were discarded, and those with diminishing quality toward their ends were truncated. These filtering steps significantly reduced the number of reads, but the retained data was of much higher quality, enhancing analytical accuracy.

Once quality filtering was completed, the forward and reverse reads were merged to reconstruct longer, contiguous sequences. This merging relies on sufficient overlap between paired reads, ensuring that the resulting sequences are both accurate and

representative of the original amplicons. Unique sequences from each sample were then compiled into an OTU table, which records the abundance of each sequence across all samples. A critical step in this process is the identification and removal of chimeric sequences, which are artifacts generated during PCR amplification. Chimeras arise when partially extended DNA strands bind to templates from different sequences and are subsequently extended, creating artificial hybrids. Removing these sequences ensures that the dataset accurately reflects the true microbial community (Smyth et al., 2010).

After the refinement of the sequences, taxonomic assignment was performed using the MIDAS database (Microbial Database for Activated Sludge). This specialized database focuses on microorganisms found in WWTP. It provides amplicon sequence variant-resolved, full-length 16S rRNA gene references with comprehensive taxonomic classification from domain to species level, enabling precise identification of community members (Dueholm et al., 2022b, 2024).

The processed data was then integrated into a single phyloseq object, a structure that facilitates downstream ecological and statistical analyses. This object combines three key components: the OTU table, which captures the abundance of each sequence; the taxonomy table, which maps sequences to their taxonomic classifications; and data obtained from the monitoring pilot plant, such as experimental conditions or operational parameters.

To evaluate sequencing depth and ensure adequate sampling effort, rarefaction curves were generated. These curves show how species richness increases with the number of reads sequenced. When the curve plateaus, it indicates that the sampling effort has sufficiently captured the diversity within the sample, minimizing the risk of underestimating richness.

Diversity was assessed at multiple levels. Alpha diversity, which measures the diversity within individual samples, is quantified using indices such as Shannon (Shannon, 1948) and Simpson (Simpson, 1949). The Shannon index measures diversity, considering both richness and evenness of abundances, making it more sensitive to the presence of rare species. Simpson index measures the probability that two individuals randomly selected from the sample will belong to the same species. To examine differences between microbial communities, beta diversity was analyzed using Non-metric Multidimensional Scaling (NMDS). This technique visualizes the relationships between samples based on the Bray-Curtis dissimilarity index (Bray & Curtis, 1957), which calculates compositional differences by considering the relative abundances of species. This index reflects the proportion of the total species abundances in which the two plots differ (Ricotta & Podani, 2017).

Finally, the relative abundance of taxa was calculated and visualized to explore community composition. Bar plots are often used to represent the proportions of different taxa across samples. By summarizing data at various taxonomic levels, such as phylum, genus, or species, these visualizations provide insights into the dominant members of the community and their shifts under different conditions.

3. RESULTS

3.1 FISH characterization

The six FISH preparations were observed under the CLSM. For quantitative FISH, a general probe set targeting all bacteria is used in combination with probes targeting the taxa of interest. Only the A sample types (unextended drops of drops of 20 µ directly placed on the slildel) were used for quantification. The A preparations were thicker and, thus, better for quantification since the fluoresce signal was stronger, and possible bias for single cell falling was lower (Daims et al., 2019). The results obtained after image processing and abundance quantification for both samples are shown in table 5. The detection of the genus-specific probe was positive in both samples. An example image is illustrated in figure 7, where it can be identified the presence of both *Thiothrix* and *Ca*. Accumulibacter.

Azonexus was not clearly detected. The fluorescence signal of its specific probe presented interferences and had a background signal. This signal was eliminated with image processing. In any case, *Azonexus* exhibited low abundances in both samples. *Thiothrix* abundances were significantly different, being more abundant in C800-SSSF1. As for *Ca*. Accumulibacter, it seems to be a dominant genus in both microbial communities.

Table 5. Relative abundances of target organisms (FISH-technique)

	Genus relative abundance (%)		
Sample	*Ca*. Accumulibacter	*Thiothrix*	*Azonexus*
SSSF-2,5	28.8	0.1	0.3
C800-SSSF1	38.5	5.9	0.7

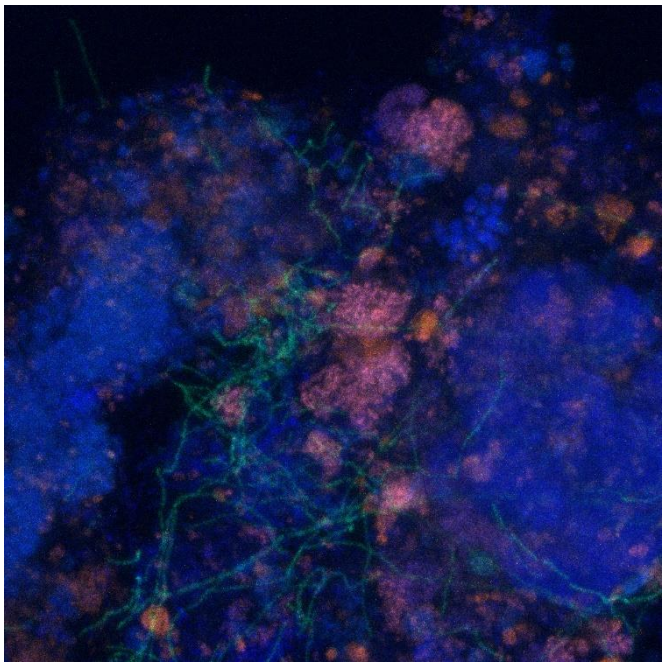

Figure 7. Z-stack image of C800-SSSF1 sample. All bacteria are depicted in blue, Ca. Accumulibacter in orange and Thiothrix in green (no presence of Azonexus in this plane).

3.2 Illumina

This section describes the results from the Illumina analysis performed by SGiEB. The bar plots in figure 8 shows the relative abundance of the genera with a relative abundance higher than 1%.

The sample C800-SSSF1 is dominated by *Thiothrix* (8,27%), the already mentioned filamentous putative PAO genus. The genus *Defluviicoccus* is also abundant (6%). It is considered a GAO (Wong & Liu, 2007) and they are capable to reduce nitrate to nitrite (Bessarab et al., 2022) and therefore assist other denitrifying populations (Maszenan et al., 2022). *Azospira* is related to N metabolism, it can use nitrate as electron acceptor. *Nitrospira* are the most common and abundant NOB in wastewater treatment systems (Daims et al., 2001) *Thauera* is abundant in denitrifying activated sludges (Morgan-Sagastume et al., 2008) and can use nitrate, nitrite or oxygen as electron acceptors (Thomsen et al., 2007). The genus *Halomonas* has been described as a putative PAO with a similar physiology to the *Accumulibacter (*Nguyen et al., 2012). The also described PAO *Dechloromonas (*now *Azonexus)* is also present with 1.82% of abundance (*Azoneus* appears as a different taxa with 0.82%, making it a 2.64%). The genera *Allochromatium, Litorilinea, Ohtaekwangia, Thiocapsa, Prosthecobacter, Rhodopirellula, Luteibacter, Chiayiivirga* and *Gimesia* have not been described as relevant for wastewater treatment and their role in the systems haven't been assessed yet (Dueholm et al., 2024).

The sample SSSF-2,5 is dominated by *Chiayiivirga* (25.88%) a relatively new genus found in WWTP (Dueholm et al., 2022a). Neither *Chiayiivirga* nor *Parasegetibacter* have their metabolism described. *Azospira* (4.80%) and *Nitrospira* (3.33%) are also present in this sample. *Bacteroides* are anaerobic fermentative organisms abundant in the gut and are during the activated sludge process is shown to decrease its presence (Niestępski et

al., 2020). *Propionivibrio* has an important role in wastewater systems, with a large present in EBPR systems, some species of the genus are considered novel GAO (Albertsen et al., 2016b). *Labrenzia, Aquabacterium, Francisella, Rubrivivax* have not been described as relevant for wastewater treatment and their role in the systems haven't been assessed (Dueholm et al., 2024). Neither Pseudacidovorax or Ohtaekwangia appear in MIDAS data base.

As for PAO microorganisms, the Illumina sequencing shows the presence of *Azonexus* (and *Dechloromonas*, now considered *Azonexus*) with a 2.67% of relative abundance. Thiothrix, is also present with a 0.34% of relative abundance.

Surprisingly, the results from Illumina analysis performed by SGiEB showed no presence of *Ca*. Accumulibacter, the already described PAO model organism. The other common PAO, *Tetrasphaera*, was not observed either. Apart from this absence, many taxa were labelled as unidentified, which was undesired. This highlights the importance of using a suitable database. In the light of these results, a reanalysis of the raw data was performed using the MiDAS database, as described in the previous section. The MiDAS database was developed specifically from samples obtained in WWTP and has become the preferred database in many wastewater studies since its development. It serves as the reference database of microbes from WWTP across the world, covering a wider taxa variability present in WWTP and including uncultured environmental taxa (Dueholm et al., 2022a).

Figure 8. Relative genus abundance (Illumina technique)

3.3 Illumina reanalysis

The results obtained from the Illumina sequencing reanalysis differ from the ones obtained from SGiEB. The same samples were analyzed and obtained relative abundance plots for the ones with abundances larger than 1% (Fig. 9).

In the case of C800-SSSF1, *Ca*. Competibacter was the most abundant genus (13.15%) and has a behavior consistent with the GAO phenotype (Kong et al., 2006). *Thiothrix* is the second most abundant genus (6.96%), followed again by *Defluviicoccus* (4.04%). *Azospira* (2.66%) and *Nitrospira* (1.45%) are also detected. Neither of the other most abundant genera (presented in figure 9) match with the ones presented in figure 8. *Azonexus* (1,12%) was also detected in the original analysis. *Ca*. Accumulibacter was detected with a low abundance 0.25%.

As for SSSF-2,5d according to the reanalysis the sample is dominated by *Ca.* Contendobacter (33.63%), part from the originally described genus *Ca.* Competibacter, this genera's behavior is consistent with the GAO phenotype (McIlroy et al., 2014). This finding is more consistent than the *Chiayiivirga* dominance reported in the SGiEB results.

Azospira (7.41%), *Neochlamydia* (2,28%) and *Nistrospira* (1.57%) are also detected in this analysis. Neither of the other most abundant genera are coincident. *Azonexus* was detected with 0.72% of abundance, similar to *Ca*. Accumulibacter with a 0.72%. *Thiothrix* presence is much lower than in the other sample, with just 0,72% detected.

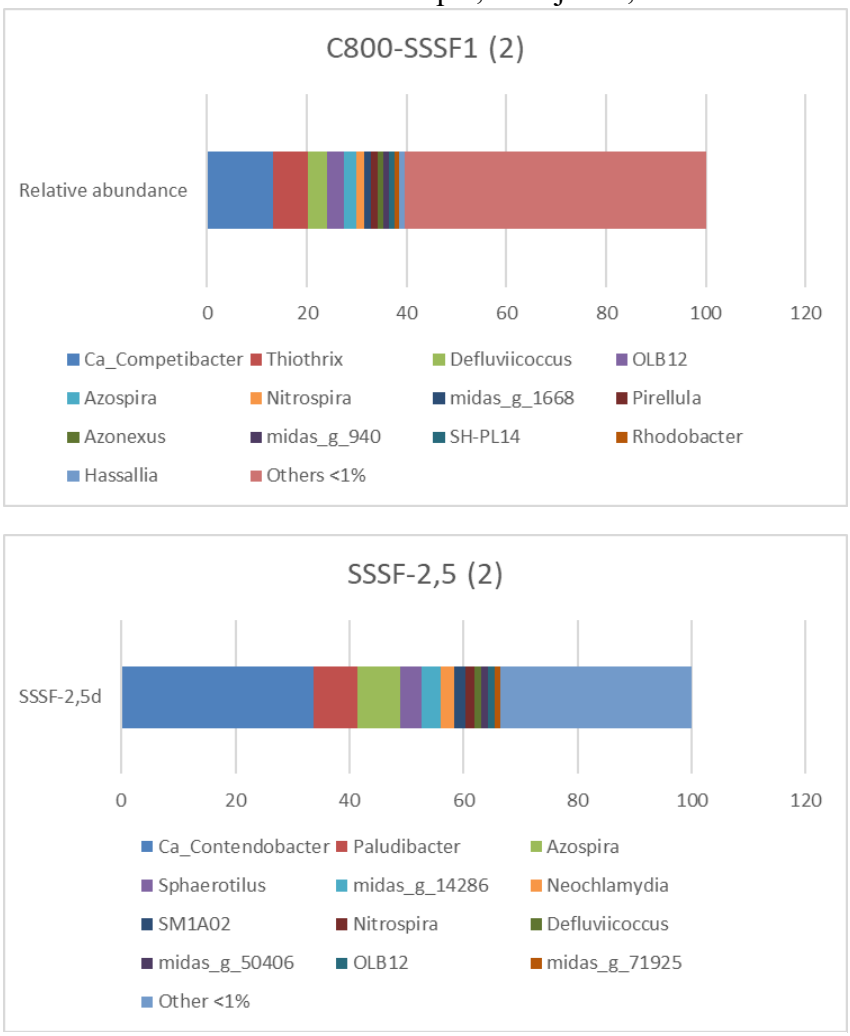

Figure 9. Relative genus abundance (Illumina reanalysis)

From the rarefaction curves obtained (Fig. 10), we can assume that the sequencing depth was sufficient to capture most of the microbial diversity in the samples since they reach a plateau. The samples with higher curves have greater species richness: C800-AER2 and C800-SSSF1. The samples with higher sample size are the ones of the system fed with real wastewater.

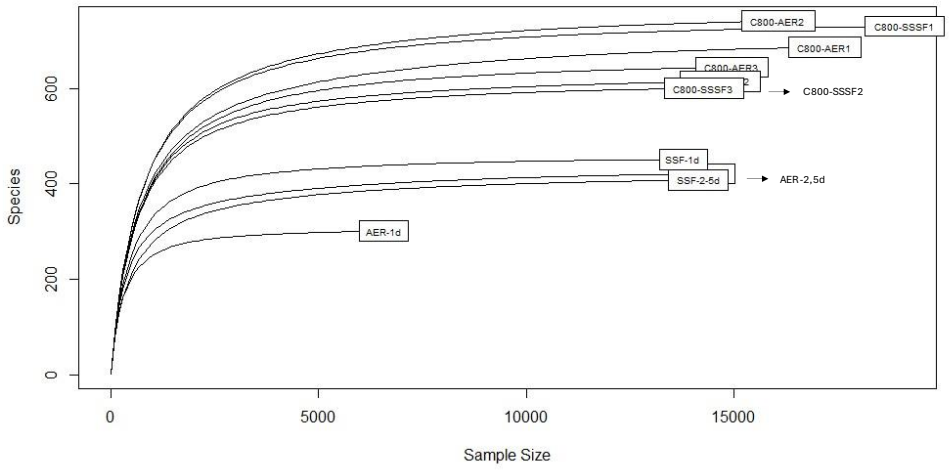

Figure 10. Rarefaction curves representing the species richness as a function of sequencing depth for different samples. The X-axis shows the sample size (number of sequences), and the Y-axis indicates the number of observed species (OTUs).

From the alpha diversity across samples the most obvious result is that both Shannon and Simpson indices consistently indicate higher microbial diversity in the samples obtained from reactors operated with real wastewater (Fig. 11). This finding suggests that real wastewater provides a more complex and diverse range of substrates, which can foster the growth of a wider variety of taxa and enhance microbial niche diversity. This trend is supported by the rarefaction curves. This result is expected, as real wastewater typically contains a heterogeneous mix of organic and inorganic compounds, trace elements and microbial presence from various sources.

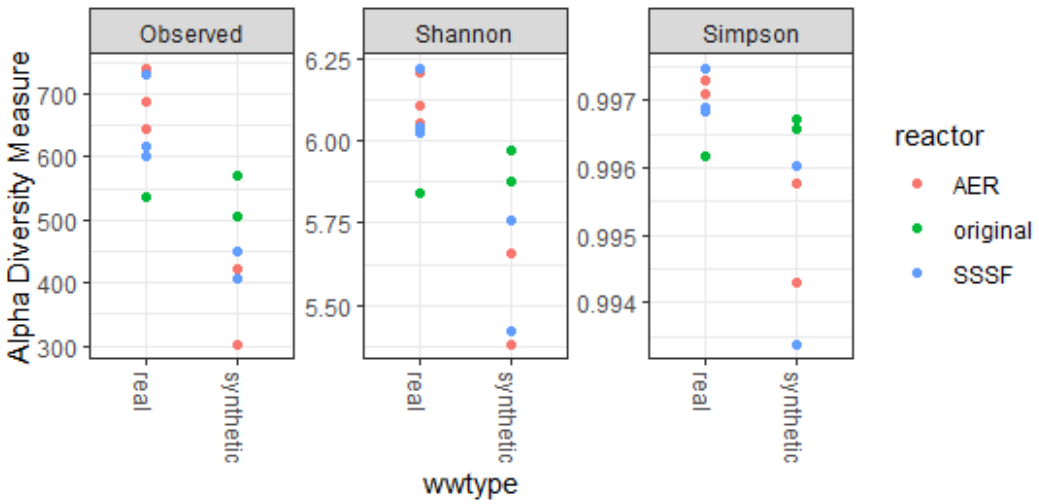

Figure 11. Alpha diversity across samples, AER and SSSF are samples from the S2EBPR plant, original is from the conventional A^2/O.

The non-metric multidimensional scaling (NMDS) analysis was used to evaluate the influence of different parameters on bacterial community composition (Fig. 12). From the monitoring parameters of the pilot plant (presented in table 1), the following were selected: reactor sampled (aerobic or SSSF, both from the S2EBPR plant), wastewater type (real or synthetic), COD in the influent and phosphorus concentration in the sampled reactor. The polygons in the NMDS plot connect groups with a similar variable value: non-continuous variables have defined polygons.

The results show no significant differences between bacterial communities across reactor type. The biomass is recirculated between reactors, which prevents the development of independent microbial communities. Although the SSSF was operated with a longer retention time (HRT=2.5d for most samples, as shown in table 1) compared to the mainstream (HRT=1d), this difference does not appear enough to modulate the bacterial community. It is likely that a larger increase in HRT in the SSSF could be a factor modulating the development of distinct microbial communities (but this is not proved in this study). In contrast, there are clear dissimilarities in relation to the type of wastewater, suggesting that the source of wastewater influences the microbial community composition.

When analyzing another important parameter, i.e. influent COD, we observe that samples with high COD levels (red) are separated from those with lower COD (blue). This suggests that COD concentration has an influence on the microbial composition. There is not an aggrupation pattern based on the phosphorus level. P concentration in the sampled reactor doesn't seem to generate differences between bacterial communities in this study. Again, it is plausible that the residence time is not different enough to modulate the bacterial composition and create distinct communities between reactors.

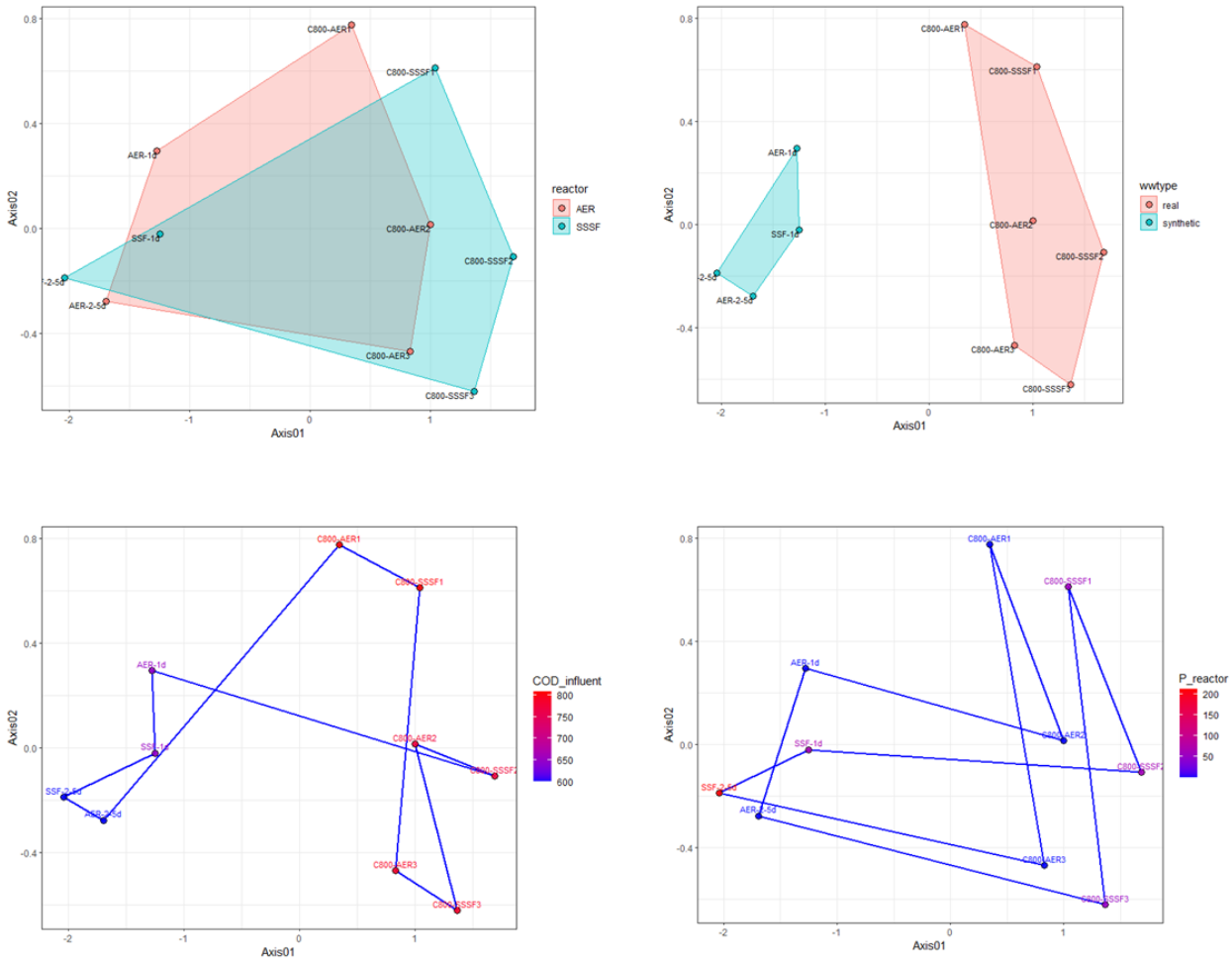

Figure 12. NDMS plots for selected monitoring parametes: reactor type, wwtype, COD concentration in the influent, P concentration in the sampled reactor.

4. DISCUSSION

The results from the two rRNA 16s analyses performed differed significantly, mainly due to the database employed (differences in the top 20 most abundant genus across samples are shown in figure 13). In wastewater research, the most used database is MiDAS (Dueholm et al., 2022a). This is the reference database because it was specifically developed using data from WWTP. In contrast, the RefSeq RDP database is a more generic database and therefore lacks coverage of many organisms that are critical to wastewater treatment process. Moreover, MiDAS has a more up-to-date and dynamic taxonomy of wastewater related microorganisms, which is continuously undergoing reclassifications. A key factor contributing to the observed differences is the variability in the reference sequences for the 16S rRNA gene across databases. If a genus is represented by only a limited number of sequences in a database, the classification algorithm may misassign reads to an incorrect taxon due to insufficient reference diversity

The bioinformatic methods used can also introduce variability. When creating the analysis pipelines, quality control filters are applied to the raw sequence data. These filtering steps can impact the pool of retained sequences and consequently alter the downstream taxonomic assignment and abundance estimations. To control and minimize this

variability, it is essential to standardize the bioinformatic workflow by using consistent quality thresholds, trimming parameters, and taxonomic databases across all samples. It is also important to use the same pipeline when the objective is to compare datasets, reducing bias introduced by methodological differences.

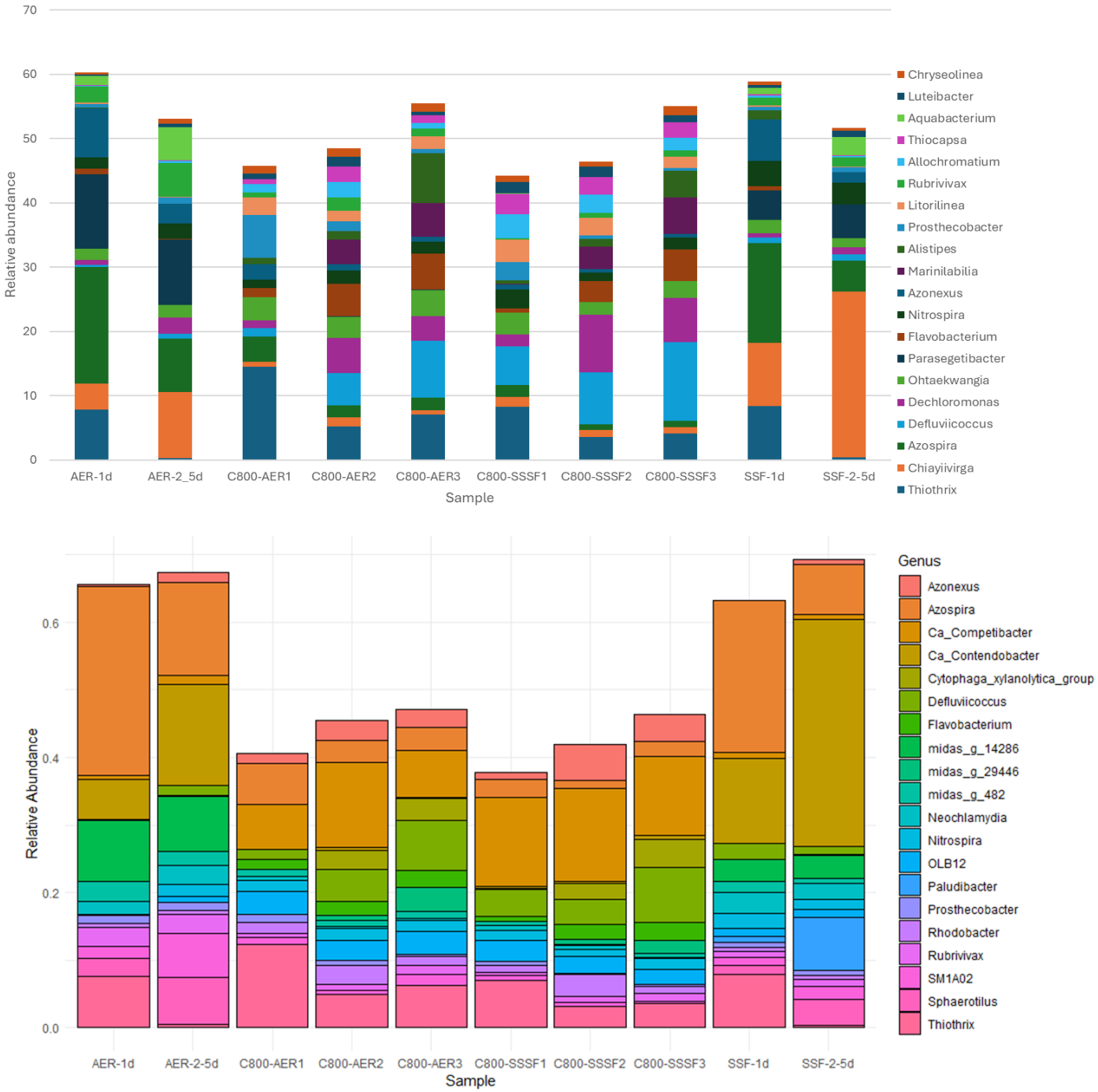

Figure 13. Bar plots showing the comparison between the TOP 20 most abundant genus of the samples; obtained by the SGiEB (top) and obtained with the data reanalysis (bottom).

There is not a strong distribution pattern from the PAO phenotype abundances across samples (Fig. 14). The only significant difference is the diminution of the putative PAO

Thiothrix with the increase of HRT. According to the results obtained *Thiothrix*'s abundance goes from 7.62% and 7.88% in AER-1d and SSSF-1d respectively, to 0.47% and 0.41% in AER-2,5d and SSSF-2,5d respectively. This result is coincident with Valverde-Pérez et al. findings.

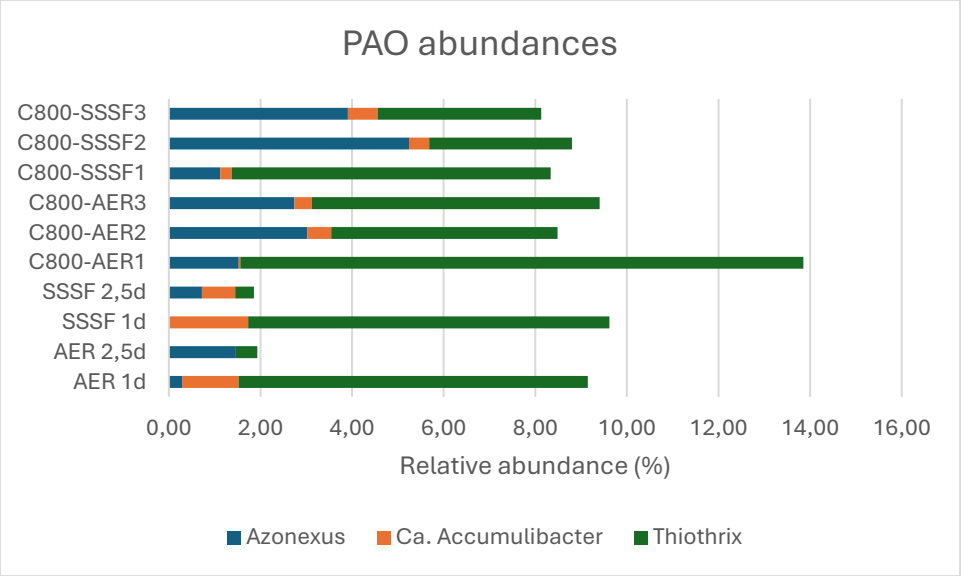

Figure 14. PAO phenotype abundances across samples

The overall look of the samples revealed some interesting insights. *Azospira* is largely abundant in synthetic wastewater samples (ranging from 7.42 to 28%) whilst on real wastewater samples it does not exceed the 6% in any of the samples. A similar pattern is found for *Ca*_Contendobacter. Meanwhile the other dominant GAO has an opposite distribution being dominant in the real wastewater samples. These differences between the microbial composition should be considered, as laboratory conditions often fail to replicate the heterogeneity and variability of real wastewater, potentially influencing the microbial community and dynamics. Thus, relying exclusively on synthetic wastewater experiments might lead to misinterpretations when extrapolating findings to full-scale wastewater treatment processes.

According to table 1, the EBPR process had good performance and the P concentration in the effluent was less than 0.20 mg/L in all the samples, indicating effective P removal. However, it is complex to make statements on the identity and contribution of PAO taxa. The results of the taxa abundances obtained with the two methods are not coincident and need to be discussed. *Ca*. Accumulibacter, *Azonexus* (former *Ca*. Dechloromonas), and *Tetrasphaera* are considered the most important PAOs, primarily due to their high frequencies and abundance in a global sampling of WWTPs (Dueholm et al., 2022a; Nielsen et al., 2019b). However, nor *Tetraspahera* (or *Ca.* Phosphoribacter) were detected and *Ca*. Accumulibacter is the least abundant PAO in all samples (according to rRNA 16S analysis).

In this study, *Ca*. Accumulibacter had a median read abundance of 0.25% and 0.72% for C800-SSSF1 and SSSF-2,5 respectively (for Illumina analysis). An extensive study to

assess PAO involved in EBPR systems (Stokholm-Bjerregaard et al., 2017b), found that *Ca*. Accumulibacter abundance ranged between 0.30 and 1.10% (with a 0.50% mean), which is consistent with our samples, with a 0.49% mean. However, this contrasts with previous FISH-based studies, such as Mielczarek et al., (2013), which evaluated the same Danish plants and reported abundances ranging from 3–8%. Similarly, Stokholm-Bjerregaard et al., (2017b) and; Ziliani et al., (2023) also observed higher values using FISH. When comparing the two methodologies used in this study, the pattern coincides, Ca. Accumulibacter abundances are much higher for FISH-quantification (28.8 and 38.5%). Rubio-Rincón et al., (2019) recently obtained large discrepancies on the genus using both methods, reporting a 60% bioabundance for FISH, while for 16s rRNA it was undetected, emphasizing the need of technique validation. A suitable explanation could be the lack of probe specificity, that remains present even though not using PAO mix as suggested (Albertsen et al., 2016a). In addition, as previously described, FISH detects rRNA 16S; larger bacteria contain more ribosomes, which results into a stronger fluorescence signal and, consequently, higher abundances in quantitative FISH. Therefore, for *Ca*. Accumulibacter, which presents cell volumes larger than the average (9.20 ± 2.4 μm3) (Fernando et al., 2019), and so larger ribosome number, it could end up overestimating its presence. Petriglieri et al., (2022) also reported differences (minor than the presented in this study) between the two methods and attributed it to the extraction efficiency, primer bias and 16S rRNA gene copy number and cell size. *Ca*. Accumulibacter has between 1-2 copies of the 16S rRNA gene, which could lead to an underestimation in Illumina results.

The former *Dechloromonas* was also assessed in the same extensive study, with an average of 1.7% abundance (Stokholm-Bjerregaard et al., 2017b)., which is consistent with our sample mean of 2%. Contrary to the situation with the previous PAO, this genus appears to be slightly overestimated by amplicon sequencing (Ahn et al., 2007; Mcilroy et al., 2011; Stokholm-Bjerregaard et al., 2017b). Our results reflect this tendency with a 0.30 and 0.70% abundance (table 5).

Thiotrix can be easily distinguished in the images due to its filamentous nature. Its size is variable between 0.60-2.50 μm. We can observe the characteristic microtubular sheath covering a line of cells (Kawasaki et al., 2017). The results obtained are the most consistent with Illumina analysis. Besides better probe specificity, software segmentation could have worked better since the nature of its morphology was outstanding.

Another bias introduced in this study is that biomass fixation was only performed for gram negative bacteria. Although the majority of bacteria in wastewater treatment are gram negative (and thus include the targeted genera), groups such as Actinobacteria or Firmicutes are gram positive, and consequently, not effectively fixed, leading to underrepresentation in the FISH quantification. As a result, the relative abundances of the targeted genera obtained may be biased. The traditional FISH technique has been proven by many authors to have limitations and biased results, as presented also in this study. These limitations include limited phylogenetic resolution difficulty to visualize certain taxa (for example populations with low ribosome number) (Engelberts et al.,

2024) and, in many occasions, low specificity probes, among others. Petriglieri, et al. (2022), highlighted the need to reevaluate and develop more specific *Ca*. Accumulibacter probes. However, in addition to the other bottlenecks, the design and optimization of new probes is time consuming and difficult to cope with the huge microbial diversity present in wastewater.

CARD-FISH is an interesting and widely adopted technique to address fluorescence signal intensity limitations (Pernthaler et al., 2002; Wilhartitz et al., 2007). Probes labelled with horseradish peroxidase (HRP) catalyze the deposition of fluorescently labelled tyramides in the target cell, increasing sensitivity up to 41 times over traditional FISH (Hoshino et al., 2008). But this technique does not resolve the low specificity problem and the labor-intensive nature of FISH methods.

An interesting alternative is GenomeFISH, genome-based fluorescence in situ hybridization. This technique uses single cell genomes obtained from environmental and clinical samples, which are amplified, fragmented, labelled, and hybridized to visualize the target microorganism *in situ.* By changing the target site it overcomes the limitations of traditional FISH and increases sensitivity and specificity (Engelberts et al., 2024).

5. CONCLUSIONS

Accurate identification is crucial for understanding microbial ecology, dynamics, and interactions in WWTPs. The high heterogeneity of the samples introduces several critical steps that can influence the final results of both techniques. Due to the diverse bacterial composition, optimizing a single protocol that performs equally well for all bacterial phyla is challenging, as the groups may respond differently to procedural step.

From this study we can conclude FISH combined with advanced techniques (e.g., genomics, proteomics) is essential for studying microbial populations and their roles. Both techniques have drawbacks and the combination of both provides complementary information. For quantification purposes, Illumina sequencing appears to be more reliable as it has less critical steps. Development of more specific FISH probes is needed for quantification purposes. Exploring alternative binding sites could also be an alternative approach for future studies.

In addition, wastewater type seems to be the most influent factor in shaping microbial communities among the parameters analyzed in this work.

6. BIBLIOGRAPHY

Ahn, J., Schroeder, S., Beer, M., McIlroy, S., Bayly, R. C., May, J. W., Vasiliadis, G., & Seviour, R. J. (2007). Ecology of the microbial community removing phosphate from wastewater under continuously aerobic conditions in a sequencing batch reactor. *Applied and Environmental Microbiology*, *73*(7), 2257–2270. https://doi.org/10.1128/AEM.02080-06/ASSET/910E6AC2-6ECD-4A06-A727-E809A7ED4FD8/ASSETS/GRAPHIC/ZAM0070776770007.JPEG

Albertsen, M., McIlroy, S. J., Stokholm-Bjerregaard, M., Karst, S. M., & Nielsen, P. H. (2016a). "Candidatus Propionivibrio aalborgensis": A novel glycogen accumulating organism abundant in full-scale enhanced biological phosphorus removal plants. *Frontiers in Microbiology*, *7*(JUL), 191473. https://doi.org/10.3389/FMICB.2016.01033/BIBTEX

Albertsen, M., McIlroy, S. J., Stokholm-Bjerregaard, M., Karst, S. M., & Nielsen, P. H. (2016b). "Candidatus Propionivibrio aalborgensis": A novel glycogen accumulating organism abundant in full-scale enhanced biological phosphorus removal plants. *Frontiers in Microbiology*, *7*(JUL). https://doi.org/10.3389/FMICB.2016.01033

Barnard, J. L. (1975). Biological nutrient removal without the addition of chemicals. *Water Research*, *9*(5–6), 485–490. https://doi.org/10.1016/0043-1354(75)90072-X

Bessarab, I., Maszenan, A. M., Haryono, M. A. S., Arumugam, K., Saw, N. M. M. T., Seviour, R. J., & Williams, R. B. H. (2022). Comparative Genomics of Members of the Genus Defluviicoccus With Insights Into Their Ecophysiological Importance. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/FMICB.2022.834906

Bray, J. R., & Curtis, J. T. (1957). An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs*, *27*(4), 325–349. https://doi.org/10.2307/1942268

Cordell, D., Drangert, J. O., & White, S. (2009). The story of phosphorus: Global food security and food for thought. *Global Environmental Change*, *19*(2), 292–305. https://doi.org/10.1016/J.GLOENVCHA.2008.10.009

Crocetti, G. R., Hugenholtz, P., Bond, P. L., Schuler, A., Keller, J., Jenkins, D., & Blackall, L. L. (2000). Identification of polyphosphate-accumulating organisms and design of 16S rRNA-directed probes for their detection and quantitation. *Applied and Environmental Microbiology*, *66*(3), 1175–1182. https://doi.org/10.1128/AEM.66.3.1175-1182.2000/ASSET/17D1E2B2-7FA7-4166-9536-0ED2FB8E2E95/ASSETS/GRAPHIC/AM0301284004.JPEG

Cyprowski, M., Stobnicka-Kupiec, A., Ławniczek-Wałczyk, A., Bakal-Kijek, A., Gołofit-Szymczak, M., & Górny, R. L. (2018). Anaerobic bacteria in wastewater treatment plant. *International Archives of Occupational and Environmental Health*, *91*(5), 571–579. https://doi.org/10.1007/S00420-018-1307-6

Daims, H., Lücker, S., & Wagner, M. (2006). daime, a novel image analysis program for microbial ecology and biofilm research. *Environmental Microbiology*, *8*(2), 200–213. https://doi.org/10.1111/J.1462-2920.2005.00880.X

Daims, H., Nielsen, J. L., Nielsen, P. H., Schleifer, K. H., & Wagner, M. (2001). In situ characterization of Nitrospira-like nitrite-oxidizing bacteria active in wastewater treatment plants. *Applied and Environmental Microbiology*, *67*(11), 5273–5284. https://doi.org/10.1128/AEM.67.11.5273-5284.2001

Daims, Holger., Nielsen, P. Halkjær. ., & Lemmer, Hilde. (2019). *FISH Handbook for Biological Wastewater Treatment: Identification and Quantification of Microorganisms in Activated Sludge and Biofilms*. IWA Publishing.

Dueholm, M. K. D., Andersen, K. S., Korntved, A. K. C., Rudkjøbing, V., Alves, M., Bajón-Fernández, Y., Batstone, D., Butler, C., Cruz, M. C., Davidsson, Å., Erijman, L., Holliger, C., Koch, K., Kreuzinger, N., Lee, C., Lyberatos, G., Mutnuri, S., O'Flaherty, V., Oleskowicz-Popiel, P., … Nielsen, P. H. (2024). MiDAS 5: Global diversity of bacteria and archaea in anaerobic digesters. *Nature Communications 2024 15:1*, *15*(1), 1–16. https://doi.org/10.1038/s41467-024-49641-y

Dueholm, M. K. D., Nierychlo, M., Andersen, K. S., Rudkjøbing, V., Knutsson, S., Arriaga, S., Bakke, R., Boon, N., Bux, F., Christensson, M., Chua, A. S. M., Curtis, T. P., Cytryn, E., Erijman, L., Etchebehere, C., Fatta-Kassinos, D., Frigon, D., Garcia-Chaves, M. C., Gu, A. Z., … Nielsen, P. H. (2022a). MiDAS 4: A global catalogue of full-length 16S rRNA gene sequences and taxonomy for studies of bacterial communities in wastewater treatment plants. *Nature Communications 2022 13:1*, *13*(1), 1–15. https://doi.org/10.1038/s41467-022-29438-7

Dueholm, M. K. D., Nierychlo, M., Andersen, K. S., Rudkjøbing, V., Knutsson, S., Arriaga, S., Bakke, R., Boon, N., Bux, F., Christensson, M., Chua, A. S. M., Curtis, T. P., Cytryn, E., Erijman, L., Etchebehere, C., Fatta-Kassinos, D., Frigon, D., Garcia-Chaves, M. C., Gu, A. Z., … Nielsen, P. H. (2022b). MiDAS 4: A global catalogue of full-length 16S rRNA gene sequences and taxonomy for studies of bacterial communities in wastewater treatment plants. *Nature Communications 2022 13:1*, *13*(1), 1–15. https://doi.org/10.1038/s41467-022-29438-7

Engelberts, P., Ye, J., Parks, D., McMaster, E., McInnes, A., Woodcroft, B., Volmer, J., McIlroy, S., & Tyson, G. (2024). *GenomeFISH: genome-based fluorescence in situ hybridisation for strain-level visualisation of microbial communities*. https://doi.org/10.21203/RS.3.RS-5531216/V1

Fernando, E. Y., McIlroy, S. J., Nierychlo, M., Herbst, F. A., Petriglieri, F., Schmid, M. C., Wagner, M., Nielsen, J. L., & Nielsen, P. H. (2019). Resolving the individual contribution of key microbial populations to enhanced biological phosphorus removal with Raman–FISH. *ISME Journal*, *13*(8), 1933–1946. https://doi.org/10.1038/s41396-019-0399-7

Hoshino, T., Yilmaz, L. S., Noguera, D. R., Daims, H., & Wagner, M. (2008). Quantification of target molecules needed to detect microorganisms by fluorescence in situ hybridization (FISH) and catalyzed reporter deposition-FISH. *Applied and Environmental Microbiology*, *74*(16), 5068–5077. https://doi.org/10.1128/AEM.00208-08/ASSET/A09F0AD0-0DC7-4FED-B095-7A0FA6D916A3/ASSETS/GRAPHIC/ZAM0160891060006.JPEG

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, *82*(11), 801–811. https://doi.org/10.1016/j.humimm.2021.02.012

Illakwahhi, D. T., Vegi, M. R., & Srivastava, B. B. L. (2024). Phosphorus' future insecurity, the horror of depletion, and sustainability measures. *International Journal of Environmental Science and Technology*, *21*(14), 9265–9280. https://doi.org/10.1007/S13762-024-05664-Y/FIGURES/5

Kanagawa, T., Kamagata, Y., Aruga, S., Kohno, T., Horn, M., & Wagner, M. (2000). Phylogenetic analysis of and oligonucleotide probe development for eikelboom type 021N filamentous bacteria isolated from bulking activated sludge. *Applied and Environmental Microbiology*, *66*(11), 5043–5052. https://doi.org/10.1128/AEM.66.11.5043-5052.2000

Kawasaki, Y., Endo, T., Fujiwara, A., Kondo, K., Katahira, M., Nittami, T., Sato, M., & Takeda, M. (2017). Elongation pattern and fine structure of the sheaths formed by Thiothrix nivea and Thiothrix fructosivorans. *International Journal of Biological Macromolecules*, *95*, 1280–1288. https://doi.org/10.1016/J.IJBIOMAC.2016.11.025

Kong, Y., Xia, Y., Nielsen, J. L., & Nielsen, P. H. (2006). Ecophysiology of a group of uncultured Gammaproteobacterial glycogen-accumulating organisms in full-scale enhanced biological phosphorus removal wastewater treatment plants. *Environmental Microbiology*, *8*(3), 479–489. https://doi.org/10.1111/J.1462-2920.2005.00914.X

Kong, Y., Xia, Y., Nielsen, J. L., & Nielsen, P. H. (2007). Structure and function of the microbial community in a full-scale enhanced biological phosphorus removal plant. *Microbiology*, *153*(12), 4061–4073. https://doi.org/10.1099/MIC.0.2007/007245-0

Kulakovskaya, T. V., Vagabov, V. M., & Kulaev, I. S. (2012). Inorganic polyphosphate in industry, agriculture and medicine: Modern state and outlook. *Process Biochemistry*, *47*(1), 1–10. https://doi.org/10.1016/J.PROCBIO.2011.10.028

Liu, X., Sheng, H., Jiang, S., Yuan, Z., Zhang, C., & Elser, J. J. (2016). Intensification of phosphorus cycling in China since the 1600s. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(10), 2609–2614.

https://doi.org/10.1073/PNAS.1519554113/SUPPL_FILE/PNAS.1519554113.SD02.XLSX

Lopez-Vazquez, C. M., Oehmen, A., Hooijmans, C. M., Brdjanovic, D., Gijzen, H. J., Yuan, Z., & van Loosdrecht, M. C. M. (2009). Modeling the PAO–GAO competition: Effects of carbon source, pH and temperature. *Water Research*, *43*(2), 450–462. https://doi.org/10.1016/J.WATRES.2008.10.032

Maszenan, A. M., Bessarab, I., Williams, R. B. H., Petrovski, S., & Seviour, R. J. (2022). The phylogeny, ecology and ecophysiology of the glycogen accumulating organism (GAO) Defluviicoccus in wastewater treatment plants. *Water Research*, *221*. https://doi.org/10.1016/J.WATRES.2022.118729

McIlroy, S. J., Albertsen, M., Andresen, E. K., Saunders, A. M., Kristiansen, R., Stokholm-Bjerregaard, M., Nielsen, K. L., & Nielsen, P. H. (2014). 'Candidatus Competibacter'-lineage genomes retrieved from metagenomes reveal functional metabolic diversity. *ISME Journal*, *8*(3), 613–624. https://doi.org/10.1038/ISMEJ.2013.162

Mcilroy, S. J., Starnawska, A., Starnawski, P., Saunders, A. M., Nierychlo, M., Nielsen, P. H., & Nielsen, J. L. (2016). Identification of active denitrifiers in full-scale nutrient removal wastewater treatment systems. *Environmental Microbiology*, *18*(1), 50–64. https://doi.org/10.1111/1462-2920.12614

Mcilroy, S. J., Tillett, D., Petrovski, S., & Seviour, R. J. (2011). Non-target sites with single nucleotide insertions or deletions are frequently found in 16S rRNA sequences and can lead to false positives in fluorescence in situ hybridization (FISH). *Environmental Microbiology*, *13*(1), 33–47. https://doi.org/10.1111/J.1462-2920.2010.02306.X

Mielczarek, A. T., Nguyen, H. T. T., Nielsen, J. L., & Nielsen, P. H. (2013). Population dynamics of bacteria involved in enhanced biological phosphorus removal in Danish wastewater treatment plants. *Water Research*, *47*(4), 1529–1544. https://doi.org/10.1016/J.WATRES.2012.12.003

Mihelcic, J. R., Fry, L. M., & Shaw, R. (2011). Global potential of phosphorus recovery from human urine and feces. *Chemosphere*, *84*(6), 832–839. https://doi.org/10.1016/j.chemosphere.2011.02.046

Mino, T., Liu, W. T., Kurisu, F., & Matsuo, T. (1995). Modelling glycogen storage and denitrification capability of microorganisms in enhanced biological phosphate removal processes. *Water Science and Technology*, *31*(2), 25–34. https://doi.org/10.1016/0273-1223(95)00177-O

Morgan-Sagastume, F., Nielsen, J. L., & Nielsen, P. H. (2008). Substrate-dependent denitrification of abundant probe-defined denitrifying bacteria in activated sludge. *FEMS Microbiology Ecology*, *66*(2), 447–461. https://doi.org/10.1111/J.1574-6941.2008.00571.X

Nagy, J., Mikola, A., Pradhan, S. K., & Zseni, A. (2019). The Utilization of Struvite Produced from Human Urine in Agriculture as a Natural Fertilizer: A Review. *Periodica Polytechnica Chemical Engineering*, *63*(3), 478–484. https://doi.org/10.3311/PPCH.12689

Nguyen, H. T. T., Nielsen, J. L., & Nielsen, P. H. (2012). "Candidatus Halomonas phosphatis", a novel polyphosphate-accumulating organism in full-scale enhanced biological phosphorus removal plants. *Environmental Microbiology*, *14*(10), 2826–2837. https://doi.org/10.1111/J.1462-2920.2012.02826.X

Nielsen, P. H., McIlroy, S. J., Albertsen, M., & Nierychlo, M. (2019a). Re-evaluating the microbiology of the enhanced biological phosphorus removal process. *Current Opinion in Biotechnology*, *57*, 111–118. https://doi.org/10.1016/J.COPBIO.2019.03.008

Nielsen, P. H., McIlroy, S. J., Albertsen, M., & Nierychlo, M. (2019b). Re-evaluating the microbiology of the enhanced biological phosphorus removal process. In *Current Opinion in Biotechnology* (Vol. 57, pp. 111–118). Elsevier Ltd. https://doi.org/10.1016/j.copbio.2019.03.008

Niestępski, S., Harnisz, M., Ciesielski, S., Korzeniewska, E., & Osińska, A. (2020). Environmental fate of Bacteroidetes, with particular emphasis on Bacteroides fragilis group bacteria and their specific antibiotic resistance genes, in activated sludge wastewater treatment plants. *Journal of Hazardous Materials*, *394*. https://doi.org/10.1016/J.JHAZMAT.2020.122544

Nouioui, I., Carro, L., García-López, M., Meier-Kolthoff, J. P., Woyke, T., Kyrpides, N. C., Pukall, R., Klenk, H. P., Goodfellow, M., & Göker, M. (2018). Genome-Based Taxonomic Classification of the Phylum Actinobacteria. *Frontiers in Microbiology*, *9*(AUG). https://doi.org/10.3389/FMICB.2018.02007

Oehmen, A., Lemos, P. C., Carvalho, G., Yuan, Z., Keller, J., Blackall, L. L., & Reis, M. A. M. (2007). Advances in enhanced biological phosphorus removal: From micro to macro scale. *Water Research*, *41*(11), 2271–2300. https://doi.org/10.1016/J.WATRES.2007.02.030

Pepin, C. A., & Wood, H. G. (1986). Polyphosphate glucokinase from Propionibacterium shermanii. Kinetics and demonstration that the mechanism involves both processive and nonprocessive type reactions. *Journal of Biological Chemistry*, *261*(10), 4476–4480. https://doi.org/10.1016/S0021-9258(17)38524-1

Pernthaler, A., Pernthaler, J., & Amann, R. (2002). Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Applied and Environmental Microbiology*, *68*(6), 3094–3101. https://doi.org/10.1128/AEM.68.6.3094-3101.2002/ASSET/AD57630D-E4FE-4F6D-B506-D8F5F411090F/ASSETS/GRAPHIC/AM0621927007.JPEG

Petriglieri, F., Petersen, J. F., Peces, M., Nierychlo, M., Hansen, K., Baastrand, C. E., Nielsen, U. G., Reitzel, K., & Nielsen, P. H. (2022). Quantification of Biologically and Chemically Bound Phosphorus in Activated Sludge from Full-Scale Plants with Biological P-Removal. *Environmental Science and Technology*, *56*(8), 5132–5140. https://doi.org/10.1021/acs.est.1c02642

Petriglieri, F., Singleton, C. M., Kondrotaite, Z., Dueholm, M. K. D., McDaniel, E. A., McMahon, K. D., & Nielsen, P. H. (2022). Reevaluation of the Phylogenetic Diversity and Global Distribution of the Genus " Candidatus Accumulibacter" . *MSystems*, *7*(3). https://doi.org/10.1128/msystems.00016-22

Petriglieri, F., Singleton, C., Peces, M., Petersen, J. F., Nierychlo, M., & Nielsen, P. H. (2021). "Candidatus Dechloromonas phosphoritropha" and "Ca. D. phosphorivorans", novel polyphosphate accumulating organisms abundant in wastewater treatment systems. *ISME Journal*, *15*(12), 3605–3614. https://doi.org/10.1038/s41396-021-01029-2

Rey-Martínez, N., Badia-Fabregat, M., Guisasola, A., & Baeza, J. A. (2019). Glutamate as sole carbon source for enhanced biological phosphorus removal. *Science of the Total Environment*, *657*, 1398–1408. https://doi.org/10.1016/j.scitotenv.2018.12.064

Ricotta, C., & Podani, J. (2017). On some properties of the Bray-Curtis dissimilarity and their ecological meaning. *Ecological Complexity*, *31*, 201–205. https://doi.org/10.1016/J.ECOCOM.2017.07.003

Rubio-Rincón, F. J., Welles, L., Lopez-Vazquez, C. M., Abbas, B., Van Loosdrecht, M. C. M., & Brdjanovic, D. (2019). Effect of lactate on the microbial community and process performance of an EBPR system. *Frontiers in Microbiology*, *10*(FEB), 435324. https://doi.org/10.3389/FMICB.2019.00125/BIBTEX

Rubio-Rincón, F. J., Welles, L., Lopez-Vazquez, C. M., Nierychlo, M., Abbas, B., Geleijnse, M., Nielsen, P. H., van Loosdrecht, M. C. M., & Brdjanovic, D. (2017). Long-term effects of sulphide on the enhanced biological removal of phosphorus: The symbiotic role of Thiothrix caldifontis. *Water Research*, *116*, 53–64. https://doi.org/10.1016/J.WATRES.2017.03.017

Ruiz-Haddad, L., Ali, M., Pronk, M., van Loosdrecht, M. C. M., & Saikaly, P. E. (2024). Demystifying polyphosphate-accumulating organisms relevant to wastewater treatment: A review of their phylogeny, metabolism, and detection. In *Environmental Science and Ecotechnology* (Vol. 21). Editorial Board, Research of Environmental Sciences. https://doi.org/10.1016/j.ese.2024.100387

Seviour, R. J., Mino, T., & Onuki, M. (2003). The microbiology of biological phosphorus removal in activated sludge systems. *FEMS Microbiology Reviews*, *27*(1), 99–127. https://doi.org/10.1016/S0168-6445(03)00021-4

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/J.1538-7305.1948.TB01338.X

Simpson, E. H. (1949). Measurement of diversity [16]. *Nature*, *163*(4148), 688. https://doi.org/10.1038/163688A0

Singleton, C. M., Petriglieri, F., Wasmund, K., Nierychlo, M., Kondrotaite, Z., Petersen, J. F., Peces, M., Dueholm, M. S., Wagner, M., & Nielsen, P. H. (2022). The novel genus, 'Candidatus Phosphoribacter', previously identified as Tetrasphaera, is the dominant polyphosphate accumulating lineage in EBPR wastewater treatment plants worldwide. *ISME Journal*, *16*(6), 1605–1616. https://doi.org/10.1038/s41396-022-01212-z

Smyth, R. P., Schlub, T. E., Grimm, A., Venturi, V., Chopra, A., Mallal, S., Davenport, M. P., & Mak, J. (2010). Reducing chimera formation during PCR amplification to ensure accurate genotyping. *Gene*, *469*(1–2), 45–51. https://doi.org/10.1016/J.GENE.2010.08.009

Stokholm-Bjerregaard, M., McIlroy, S. J., Nierychlo, M., Karst, S. M., Albertsen, M., & Nielsen, P. H. (2017a). A critical assessment of the microorganisms proposed to be important to enhanced biological phosphorus removal in full-scale wastewater treatment systems. *Frontiers in Microbiology*, *8*(APR), 247313. https://doi.org/10.3389/FMICB.2017.00718/BIBTEX

Stokholm-Bjerregaard, M., McIlroy, S. J., Nierychlo, M., Karst, S. M., Albertsen, M., & Nielsen, P. H. (2017b). A critical assessment of the microorganisms proposed to be important to enhanced biological phosphorus removal in full-scale wastewater treatment systems. *Frontiers in Microbiology*, *8*(APR), 247313. https://doi.org/10.3389/FMICB.2017.00718/BIBTEX

Tayà, C., Garlapati, V. K., Guisasola, A., & Baeza, J. A. (2013). The selective role of nitrite in the PAO/GAO competition. *Chemosphere*, *93*(4), 612–618. https://doi.org/10.1016/j.chemosphere.2013.06.006

Thomsen, T. R., Kong, Y., & Nielsen, P. H. (2007). Ecophysiology of abundant denitrifying bacteria in activated sludge. *FEMS Microbiology Ecology*, *60*(3), 370–382. https://doi.org/10.1111/J.1574-6941.2007.00309.X

Valverde-Pérez, B., Wágner, D. S., Lóránt, B., Gülay, A., Smets, B. F., & Plósz, B. G. (2016). Short-sludge age EBPR process – Microbial and biochemical process characterisation during reactor start-up and operation. *Water Research*, *104*, 320–329. https://doi.org/10.1016/J.WATRES.2016.08.026

Wanner, J., Kucman, K., Ottová, V., & Grau, P. (1987). Effect of anaerobic conditions on activated sludge filamentous bulking in laboratory systems. *Water Research*, *21*(12), 1541–1546. https://doi.org/10.1016/0043-1354(87)90139-4

Wilhartitz, I., Mach, R. L., Teira, E., Reinthaler, T., Herndl, G. J., & Farnleitner, A. H. (2007). Prokaryotic community analysis with CARD-FISH in comparison with FISH in ultra-oligotrophic ground- and drinking water. *Journal of Applied Microbiology*, *103*(4), 871–881. https://doi.org/10.1111/J.1365-2672.2007.03319.X

Wong, M. T., & Liu, W. T. (2007). Ecophysiology of Defluviicoccus-related tetrad-forming organisms in an anaerobic-aerobic activated sludge process. *Environmental Microbiology*, *9*(6), 1485–1496. https://doi.org/10.1111/J.1462-2920.2007.01267.X

Young, A. D., & Gillung, J. P. (2020). Phylogenomics — principles, opportunities and pitfalls of big-data phylogenetics. In *Systematic Entomology* (Vol. 45, Issue 2, pp. 225–247). Blackwell Publishing Ltd. https://doi.org/10.1111/syen.12406

Yuan, Z., Pratt, S., & Batstone, D. J. (2012). Phosphorus recovery from wastewater through microbial processes. *Current Opinion in Biotechnology*, *23*(6), 878–883. https://doi.org/10.1016/J.COPBIO.2012.08.001

Zhou, Y., Pijuan, M., Oehmen, A., & Yuan, Z. (2010). The source of reducing power in the anaerobic metabolism of polyphosphate accumulating organisms (PAOs) – a mini-review. *Water Science and Technology*, *61*(7), 1653–1662. https://doi.org/10.2166/WST.2010.983

Ziliani, A., Bovio-Winkler, P., Cabezas, A., Etchebehere, C., Garcia, H. A., López-Vázquez, C. M., Brdjanovic, D., van Loosdrecht, M. C. M., & Rubio-Rincón, F. J. (2023). Putative metabolism of Ca. Accumulibacter via the utilization of glucose. *Water Research*, *229*, 119446. https://doi.org/10.1016/J.WATRES.2022.119446