Quantitative DNA stable isotope probing identifies active microorganisms assimilating volatile fatty acids in full-scale enhanced biological phosphorus removal processes

Sampara. P.[1], Andy Tomatsu[2], Rex R. Malmstrom[2], Ziels. R. M[1,3,*]

[1] Department of Civil Engineering, The University of British Columbia, Vancouver, BC, Canada

[2] Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley CA, USA

[3] Genome Sciences and Technology Program, The University of British Columbia, Vancouver, BC, Canada

* Corresponding author

e-mail correspondence: rziels@civil.ubc.ca

Abstract

Enhanced biological phosphorus removal (EBPR) systems often rely on exogenous carbon sources like volatile fatty acids (VFAs) to achieve higher P removal. Here, we employed DNA quantitative stable isotope probing (qSIP) using two VFAs, acetate and propionate, in cyclic anaerobic/aerobic incubations to assess their effects on P cycling and microbial activity with biomass from two full-scale EBPR water resource-recovery facilities that utilize VFA addition. We found that anaerobic VFA uptake preferences differed within known groups of PAOs, such as *Candidatus* Accumulibacter and *Tetrasphaera*-affiliated members (e.g., *Ca*. Phosphoribacter), between the two biomasses. The combination of qSIP with metagenomics identified isotopically labelled phages that were linked to active PAOs, highlighting their roles in modulating EBPR community composition and activity. The highest levels of anaerobic labeling from acetate were in genomes belonging to *Saccharimonadales* and *Rickettsiales*, which are generally host-associated with bacteria and eukaryotes, respectively. This finding highlights the possibility of cross-feeding between PAO hosts and their parasites or predators, as well as the role of so-far uncharacterized organisms participating in carbon cycling under EBPR conditions. Collectively, these results expand our understanding of the ecological interactions involved in communities anaerobically uptaking VFAs and cycling P that are central to EBPR.

Keywords

Polyphosphate accumulating organisms, metagenomics, phage, wastewater treatment, nutrient removal

Synopsis

DNA stable isotope probing revealed carbon preferences in known polyphosphate accumulating organisms (PAOs), active phage that infect them, as well as potentially novel PAOs within two full-scale facilities.

Introduction

Water resource recovery facilities (WRRFs) employ enhanced biological phosphorus removal (EBPR) to sequester soluble phosphorus (P) from wastewater, safeguarding downstream environments from aquatic pollution. Traditional EBPR facilities cycle biomass between anaerobic and aerobic conditions to support the growth of polyphosphate accumulating organisms (PAOs)[1]. More P is taken up in the aerobic phase than is released in the anaerobic phase, and subsequently P is removed by wasting biomass with poly-P stores[2]. However, EBPR typically relies on the addition of exogenous volatile fatty acids (VFAs) (e.g., acetate or propionate) during the anaerobic phase to optimize PAO growth and phosphorus removal[3]. Recently, WRRF process configurations that internally generate VFAs through sludge fermentation are gaining attention for their resilience to fluctuating carbon loads[4,5]. Such process configurations maintain a consistent VFA supply, avoiding the need for exogenous carbon addition[4,6]. Optimizing fermentation operational conditions to provide favorable VFAs that promote P removal efficiency may be needed for broader adoption[6].

The impact of different VFAs (e.g., acetate vs. propionate) on P removal dynamics is currently ambiguous. Some studies reported that acetate is the optimal substrate for higher P removal[7,8], while others found propionate provided a competitive advantage to PAOs[9,10], and others showed a similar preference between acetate and propionate by PAOs[11]. These contradictory findings regarding VFA preference might be explained, at least in part, by differences in PAO population composition[12] and/or environmental conditions like temperature[13], but untargeted measurements of anaerobic substrate uptake by PAOs are sparse[14,15]. Resolving the source of this ambiguity could help to facilitate strategies to improve EBPR performance by tailoring VFA production for a given PAO community.

Additionally, an enormous amount of unexplored microbial diversity exists within WRRFs[16]. Identifying novel members, evaluating their physiological traits, and linking their identity and metabolic potential with key ecosystem functions, such as P removal, can help to expand the potential of emerging bioprocesses for resource recovery[17]. However, specifically linking taxonomic identity to *in situ* function is challenging. DNA stable isotope probing (SIP) is a powerful approach that establishes links between substrate assimilation and genomic identity by tracking growth on isotopically labelled substrates[18]. Quantitative SIP (qSIP) is a recent advancement in DNA-SIP to measure atom fraction excess (AFE), which is the increase in the isotopic labeling of DNA above background levels[19–22]. While 16S rRNA gene-based qSIP can identify *in situ* active microbes and estimate their AFE, qSIP has been recently adapted for metagenomic sequencing to establish stronger links between *in situ* active taxa, their metabolic potential, and ecosystem-specific functions[22]. In this study, we leveraged both 16S rRNA gene-based qSIP and metagenomic qSIP to investigate anaerobic assimilation of ^{13}C-acetate and ^{13}C-propionate by microbial communities from two full-scale EBPR WRRFs that generate VFAs via primary sludge fermentation. Using this approach, we identify substrate preferences of PAOs, establish linkages between PAOs and active phages, and also uncover microbes that are possible PAOs based on metabolic capacity and anaerobic carbon assimilation patterns.

Methods

Source of EBPR biomass

Activated sludge biomass was sampled from the EBPR systems at the City of Penticton and the Westside Regional WRRFs (both in British Columbia, Canada). Both wastewater treatment plants have a similar EBPR process configuration (Figure S1), with an anaerobic/anoxic/aerobic (A_2O) activated sludge configuration. Primary sludge is fermented and the VFA-rich stream is supplied to the anaerobic zone along with the return activated sludge. Mixed liquor samples were collected from the aerobic zone of the bioreactors for SIP incubations, as well as for time-series analysis of microbial community composition (Text S1). A total of eight time-series samples were collected from the bioreactor in Penticton, and two samples were collected from the Westside Regional facility. SIP tests on the freshly collected samples were initiated within 8 hours of sampling (Text S1).

Cyclic anaerobic/aerobic SIP incubations with acetate and propionate as VFAs

Prior to the start of the SIP incubation batch tests, samples were aerated for one hour to remove residual phosphorus from the wastewater matrix. After this pre-incubation phase, 100 mL biomass from either treatment plant was washed with PBS, divided into 10 mL triplicate treatments in 15 mL glass vials sealed with butyl rubber stoppers and flushed with N_2. Vials were then amended either with acetate or propionate (^{12}C or ^{13}C, 250 mg COD/L), resulting in a total of four overall treatments: 'Penticton+acetate', 'Penticton+propionate', 'Westside+acetate', and 'Westside+propionate'. The ^{13}C acetate and propionate were universally labelled (>99% ^{13}C content, Cambridge Isotopes). A control incubation without a carbon source was also included. Each SIP incubation involved five cycles consisting of a 3-hour anaerobic phase (N_2 purged, with carbon sources amended at the beginning) followed by a 2-hour aerobic phase (oxygen saturated,

with 25 mg P/L amended at the beginning). After each anaerobic or aerobic phase, samples were collected for P and COD measurements, and then biomass was washed via centrifugation (2,250*xg* for 3 minutes) followed by resuspension in phosphate-buffered saline at pH 7.5 to minimize carbon carryover. All incubations were shaken at 800 rpm to ensure ample mixing throughout all phases. Further details of the SIP incubations are provided in Text S1. DNA was extracted from the treatment replicates after the end of five anaerobic-aerobic cycles (Text S2).

Isopycnic centrifugation, gradient fractionation, and qPCR

Prior to density gradient centrifugation, the extent of isotopic enrichment of bulk DNA was confirmed with ultrahigh-performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS) (see Text S2). DNA extracts from the biological triplicates of the 'Westside+acetate' incubation were subjected to isopycnic centrifugation and gradient fractionation at the Joint Genome Institute (JGI), following the protocol of Nuccio et al.[23]. DNA was fractionated into 16 density fractions per sample, except for one replicate from the ^{12}C incubation that resulted in 15 fractions, yielding a total of 95 fractions for shotgun metagenomic sequencing. Synthetic DNA standards known as 'sequins'[24] were added to density fractions to estimate the absolute concentration of genomic features[22]. A second set of synthetic DNA standards were added to samples prior to centrifugation and fraction collection to ensure quality control of the sequencing pipeline and to aid in the detection of anomalous samples[22]. Metadata for these pre-centrifugation spike-ins is reported in Datasheet S1.

The other three SIP incubation DNA extracts underwent isopycnic centrifugation and density gradient fractionation based on the method proposed by Neufeld et al.[25], which does not utilize synthetic DNA standards (Text S2). DNA from each replicate was fractionated into 18 fractions. A replicate from the ^{13}C 'Penticton+propionate' incubation biomass was discarded due to

disturbance of the gradient during fractionation (Figure S2). Ultracentrifuged DNA in CsCl-gradient-buffer was purified using precipitation using 30% PEG6000 and GlycoBlue (ThermoFisher Scientific; AM9515; United States) as the co-precipitant. The DNA concentration of purified fractionated samples was quantified using the High Sensitivity Assay kit on a Qubit 4.0 Fluorometer (ThermoFisher Scientific, United States). The total concentration of 16S rRNA genes was estimated in each purified density fraction through qPCR assays with 515F/806R primers[26]. Metadata of each incubation, including isotope, replicate, fraction number, buoyant density, and qPCR-based estimation of 16S rRNA gene copies are provided in Datasheets S2-S5, in accordance with the Minimum Information for any Stable Isotope Probing Sequence (MISIP) reporting standards[27].

Metagenomics and 16S rRNA amplicon sequencing

Metagenomic sequencing (2x150bp paired-end sequencing) of all 95 fractions recovered from the replicates of the 'Westside+acetate' biomass was performed on the Illumina NovaSeq platform at the JGI (Text S2). Additionally, every buoyant density fraction from all SIP incubations underwent partial 16S rRNA gene amplicon sequencing on an Illumina MiSeq in 2x300bp paired-end mode (Text S2). Prior to 16S rRNA gene amplicon sequencing, the fractionated DNA was pre-amplified using non-barcoded 515F/926R primers[26] and Q5 Hot Start High-Fidelity 2X Master Mix (New England BioLabs, USA) for 18 PCR cycles (Table S1). Subsequently, library preparation was conducted with the pre-amplified product using barcoded 515F/926R primers and PhusionPlus PCR mastermix (Thermo Fisher Scientific, United States), as described by Comeau et al.[28]. A no-template control from the pre-amplification step was included in library preparation.

Unlabelled bulk DNA from the 'time-zero' (i.e. after pre-incubation) samples from both Penticton and Westside Regional WRRFs was sequenced on the PacBio Sequel IIe platform in HiFi mode at the Sequencing Center at Brigham Young University (Text S2). Lastly, time-series DNA

155 samples collected from the WRRFs were subjected to 16S rRNA gene amplicon sequencing
156 targeting the V4 region using 515F/806R primers[29] on an Illumina MiSeq in 2x300bp paired-end
157 mode (Text S2). Sequencing throughput and NCBI BioSample accessions for each metagenomic
158 library is supplied in Datasheet S6.

159 ***Metagenome assembly and binning***

160 The PacBio HiFi reads from Penticton and Westside Regional time-zero samples were individually
161 assembled using Hifiasm-meta(v0.3.1)[30,31] and mapped using minimap2(v2.24-r1122)[32] (Text S3
162 and Figure S3). Both Penticton and Westside Regional PacBio assemblies were independently
163 binned using MetaBAT2(v2.12.1)[33], MaxBin2(v2.2.4)[34], and Vamb(v4.1.3)[35] to recover
164 metagenome assembled genomes (MAGs). DASTool(v1.1.6)[36] was then applied to these MAGs
165 for consensus binning and de-replication (see Text S3 for details).
166 For SIP-metagenomics analysis of the 'Westside+acetate' incubation, quality filtered short-reads
167 from the 95 density fractions were co-assembled using MetaHipMer(v2.0.1.2)[33]. The 95 buoyant
168 density gradient metagenomes were then mapped separately to the Westside HiFi assembly and
169 the MetaHipMer co-assembly using BowTie2(v2.3.4.3)[38]. SAMtools(v1.9)[39] was used to convert
170 mapping data to BAM files. Binning was performed using MetaBAT2 and Vamb, and DASTool
171 was used to produce consensus MAGs separately from the Westside Regional HiFi assembly
172 and the MetaHipMer co-assembly. The resulting two sets of consensus MAGs were then finally
173 de-replicated using dRep(v3.4.3)[40] (MASH threshold: 0.8; ANImf threshold: 0.95) to obtain the
174 final set of MAGs. Quality filtered short-reads from all the 95 buoyant density fractions were then
175 mapped to this final set of MAGs using BowTie2 to obtain coverage values. MAG coverages are
176 provided in Datasheet S7. CheckM2(v1.0.2)[41] was used to assess MAG completion and
177 contamination, and GTDB-Tk(v1.7.0)[42] was used for taxonomic classification based on the GTDB
178 (release 214)[43] (Datasheet S8). DRAM(v.1.4.6)[44] was used to annotate MAGs.

Viral contigs were identified using VirSorter2(v2.24)[45] and quality assessment was performed using CheckV(v1.0.1)[46] (see Text S3). The phage-host associations were determined using iPHoP(v.1.3.3)[47]. Phage taxonomic classification was performed using PhaGCN[48] and phage lifestyle was predicted using PhaTYP[49] on the PhaBOX webserver[50].

16S rRNA ASV recovery

Amplicon reads from the SIP buoyant density fractions and the time-series sequencing data were quality filtered individually using cutadapt(v4.4)[51] and the DADA2 'filterAndTrim' program to ensure quality score of reads was above 30 (Text S3). Thereafter, all SIP-related samples (targeting the V3-V4 region of the 16S rRNA gene) were collectively processed using the DADA2(v1.28.0)[52] R package to obtain a single set of ASVs from the SIP incubations. Any ASV that had reads detected in the no-template control was removed from further analysis. The time-series data (targeting the V4 region of the 16S rRNA gene) was processed similarly. Taxonomic classification on both datasets was performed using the MiDAS database(v5.2)[53,54].

Estimation of per-taxon isotopic enrichment with quantitative SIP (qSIP)

Overall, isotopic enrichment of microbial populations and their atom fraction excess (AFE) was estimated based on the qSIP framework proposed by Hungate et al.[19] using inferred absolute abundance of microbial populations across the buoyant density gradient. This framework was applied for both the 16S rRNA gene amplicon sequencing and metagenomic sequencing SIP datasets.

Estimating atom fraction excess with 16S rRNA gene-based qSIP

For the 16S rRNA gene qSIP data, AFE was estimated for the filtered ASV datasets using ASV absolute abundances based on qPCR and amplicon sequencing-based relative abundance

values in every fraction, using the HTSSIP[55] R package. Absolute abundance of each ASV was estimated within each fraction as the product of relative abundance of each ASV and total copies of 16S rRNA genes measured by qPCR in that fraction. ASVs with zero abundance in more than eight fractions collectively across all replicates, as well as with a GC content estimated to be ≥99%, were filtered from downstream analysis. In the amplicon-based qSIP analysis, GC content is estimated using an empirical fit of weighted buoyant density in the unlabelled control[19].

To account for multiple hypothesis testing in AFE values, a false coverage rate (FCR) correction was implemented on the bootstrapped (10,000 times) datasets using the HTSSIP[55] R package. Isotope incorporation was considered positive if the lower 95% confidence interval (FCR-corrected) of bootstrapped AFEs was above zero.

Estimating atom fraction excess with qSIP metagenomics

For the metagenomic SIP data, the modified qSIP bootstrapping approach was used as discussed in Vyshenska et al.[22], using the *SIPmg* R package(v1.5.1). In metagenomic SIP, the GC content of the species is directly inferred from the MAG sequence, rather than empirically derived as in 16S rRNA gene amplicon-based qSIP. To determine absolute abundance of MAGs across the buoyant density gradient, a standard curve was plotted for each sample between coverage of sequins and their known added concentration into the sample. A robust linear regression was performed on sequin metagenomic coverage versus molarity using the *MASS* package[56] in R within the *SIPmg* R package. In two of the 95 metagenomes, sequins were unintentionally excluded. In lieu of the sequins in these two samples, an average of regression parameters was performed across the fractions with the closest buoyant density, across ^{12}C and ^{13}C triplicate samples. The regression parameters were then used to recover absolute concentrations of MAGs and viral contigs within those two metagenomes. Metadata of sequins is supplied in Datasheet S9 and the regression parameters for scaling coverages to absolute concentrations is provided in Datasheet S10. Quality control using the pre-centrifugation spike-ins is described in Text S4.

The inferred absolute concentration of MAGs (Datasheet S11) and viral contigs within each density fraction was then used to estimate weighted mean buoyant density as described in Vyshenska et al.[57] using the *SIPmg* R package. The equations for AFE estimation described by Hungate et al.[19] were then used to estimate AFE values for each MAG and viral contig. Similar to the 16S rRNA ASV qSIP pipeline, 10,000 bootstraps were performed for FCR correction to determine 95% confidence intervals of AFE, and isotope incorporators were inferred if the corrected lower confidence interval of AFE was positive.

Data availability

The 16S rRNA gene amplicon and metagenome sequencing data, including raw data and MAGs, were deposited in the NCBI BioProject database under the accession numbers PRJNA1122930 and PRJNA1120183, respectively. (Reviewer link for yet to be released Bioproject PRJNA1122930: https://dataview.ncbi.nlm.nih.gov/object/PRJNA1122930?reviewer=pth81512cj8e917vu4opg3oahp)

Results

Cyclic anaerobic/aerobic SIP incubations of biomass with acetate and propionate as the substrates

Cyclic anaerobic/aerobic SIP incubations were carried out by providing acetate or propionate anaerobically, and washing the biomass between redox changes to minimize aerobic carbon assimilation. Characteristic EBPR behavior, involving P release and uptake in the anaerobic and aerobic phases, respectively, was observed in all four SIP incubations (Figure 1; Table 1 and S2). Further, biomass in the control treatment without a carbon source showed minimal P cycling (Figure 1; Table S2), indicating that the anaerobic/aerobic SIP incubation conditions promoted PAO metabolism.

The P cycling behavior in the incubations indicated a preference for anaerobic uptake of different VFAs between the WRRFs. The Westside Regional WRRF biomass experienced greater anaerobic P release and uptake with acetate (mean P uptake across five phases in biological triplicate: 99.5 ± 0.28%) than propionate (mean P uptake: 52.6 ± 25.5, n=3) (Table 1 and S2). In contrast, the Penticton biomass had greater P release and uptake with propionate (mean P uptake: 93.2 ± 4.44%, n=3) versus acetate (mean P uptake: 20.5 ± 14.1, n=3) (Table 1 and S2). Despite different P cycling performance, biomasses from both WRRFs preferentially consumed propionate during the anaerobic period (mean uptake across five anaerobic phases ^{13}C samples in Penticton: 83.9 ± 4.3%, n=3; Westside Regional: 91.9 ± 8.3%, n=3) compared to acetate (Penticton: 39.4 ± 30.8%, n=3; Westside Regional: 36.9 ± 15.2%, n=3) (Table 1).

P concentrations within the ^{12}C batch tests followed similar trends as those in the ^{13}C tests (Table 1). There were no statistically significant differences (α=0.05, Welch Two Sample t-test) in P concentrations between ^{12}C or ^{13}C treatments of any carbon source in any of the incubations,

except within the aerobic phase of both the VFA incubations with Penticton biomass (Table S3). COD uptake was also not statistically significantly different (α=0.05, Welch Two Sample t-test) between isotopic treatments within any SIP incubation batch test (Table S3).

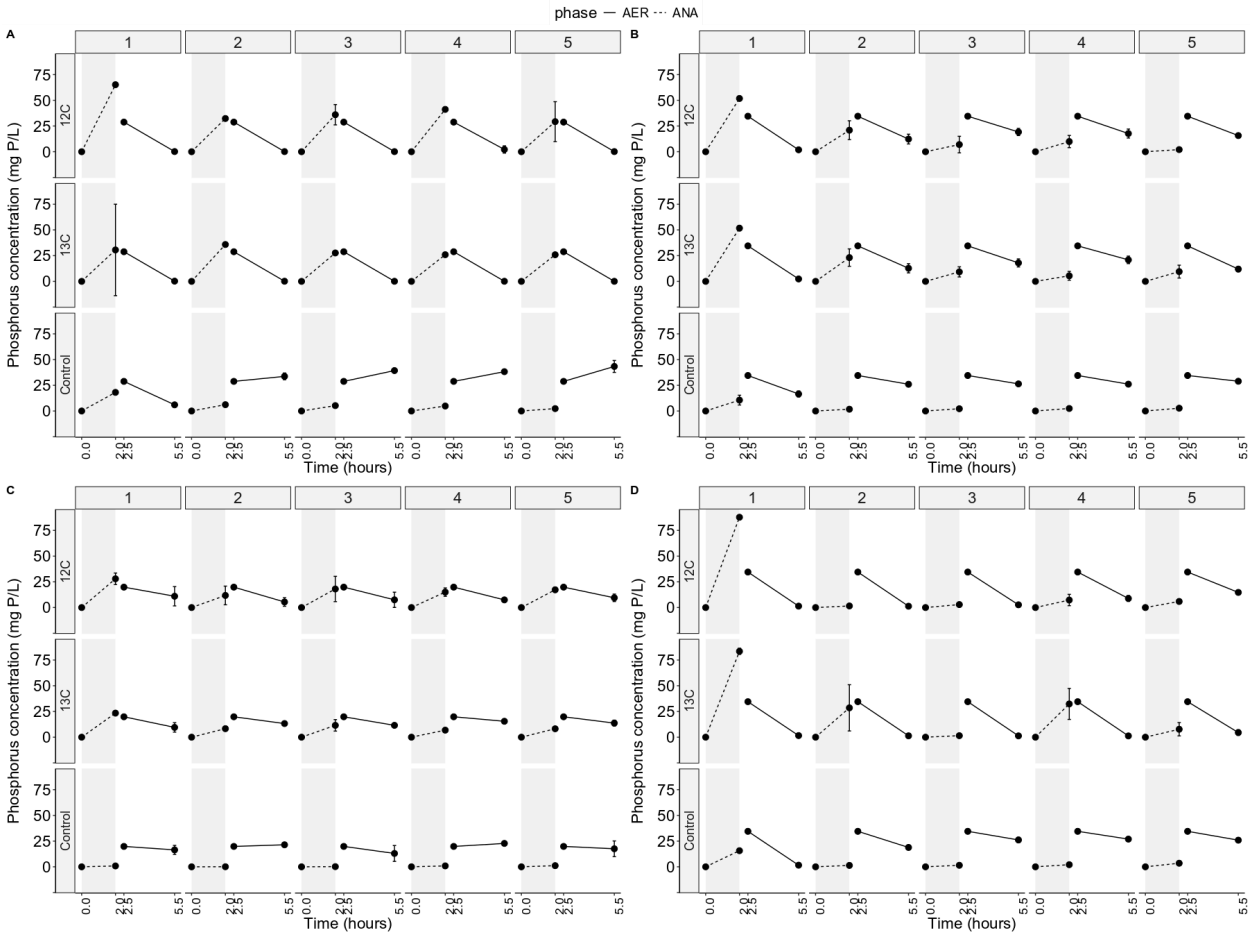


Figure 1. P dynamics in the five aerobic (light shade; 'AER') and anaerobic (dark shade; 'ANA') phases in the ^{12}C, ^{13}C, and control triplicates from the (A) 'Westside+acetate', (B) 'Westside+propionate', (C) 'Penticton+acetate', and (D) 'Penticton+propionate' SIP incubations. Half-hour gaps in the timeline indicate the wash step between phases. P concentrations in the provided medium were 0 mg/L at the start of the anaerobic phase, and 25 mg/L P at the start of the aerobic. Data points are mean values, and error bars are shown but may be shielded by certain data points.

Table 1. Mean aerobic P uptake and anaerobic COD uptake percentages across five phases in the triplicate isotopic incubations in the four SIP batch tests.

SIP Batch Test	^{12}C		^{13}C	
	Aerobic P Uptake (%)	**Anaerobic COD Uptake (%)**	**Aerobic P Uptake (%)**	**Anaerobic COD Uptake (%)**
Penticton+acetate	48.8 ± 13.3	31.7 ± 38.5	20.5 ± 14.1	39.4 ± 30.8
Penticton+propionate	78.6 ± 22.3	84.0 ± 7.9	93.2 ± 4.4	83.9 ± 4.3
Westside+acetate	97.7 ± 3.98	36.8 ± 13.2	99.5 ± 0.28	36.9 ± 15.2
Westside+propionate	52.0 ± 24.9	85.7 ± 13.1	52.6 ± 25.5	91.9 ± 8.3

Bulk DNA enrichment and microbial community composition in SIP batch tests

As expected, all biomass fed with ^{13}C substrate had a significantly higher ^{13}C enrichment of DNA compared to ^{12}C fed biomass based on ultrahigh-performance liquid chromatography-tandem mass spectrometry (α=0.05, Welch Two Sample t-test) (Table S4). Subsequent qPCR on the buoyant density fractionated samples also revealed a slight increase in DNA concentration at heavier buoyant densities among ^{13}C samples relative to the ^{12}C incubations (Figure S4). A total of 367 buoyant density fractions then underwent partial 16S rRNA gene amplicon sequencing to resolve the isotope incorporating taxa within the SIP incubations.

The microbial community composition differed between Westside and Penticon (Figure S5), and community composition changed over the course of the incubation based on the VFA substrate added (PERMANOVA with 10,000 permutations, Westside Regional WRRF biomass: R^2 = 0.06, F-statistic = 11.6, *p-value*<0.05; Penticton WRRF: R^2 = 0.03, F-statistic = 5.8, *p-value*<0.05).

Among the EBPR-relevant taxa (i.e., what is considered a known PAO or a glycogen accumulating organism (GAO)), ASVs classified to the *Tetrasphaera* genus were relatively abundant across all incubations (Figure S6), mirroring the microbial community composition observed within the time-series analysis conducted on the full-scale reactors throughout the sampling period for SIP incubations (Figure S7). Also, EBPR-relevant members, such as *Candidatus (Ca.)* Accumulibacter, *Ca*. Phosphoribacter, *Acinetobacter*, and *Ca*. Competibacter, were among the top 30 abundant taxa in one or more of these SIP incubations (Figure S6).

Quantification of DNA isotope incorporation of 16S rRNA ASVs in the SIP incubations

A total of 105 16S rRNA gene ASVs were identified as isotopic incorporators in the 'Westside+acetate' incubation, 388 in 'Westside+propionate', 84 in 'Penticton+acetate', and 725 ASVs in 'Penticton+propionate', representing 5.12%, 6.86%, 1.55%, and 14.9% of the total ASVs recovered in these communities, respectively. The higher number of metabolically active ASVs in the propionate amended incubations is consistent with the greater COD uptake observed for this substrate (Table 1). The EBPR-relevant genera of *Tetrasphaera* and *Ca.* Accumulibacter showed WRRF-specific VFA preferences by only assimilating acetate in the Westside biomass, whereas they incorporated propionate within both WRRF biomasses (Figure 2). Taxa from the EBPR-relevant genus of *Ca.* Phosphoribacter were active across all incubations (Figure 2). ASVs related to *Acinetobacter*, previously proposed as likely PAOs[58], also incorporated acetate and propionate in the Penticton incubations (Figure 2). *Dechloromonas* members were observed at a lower relative abundance (<0.5%) in the SIP incubations and bulk biomasses from the full-scale bioreactors of both facilities and the four SIP incubations, but exhibited the highest AFE among the observed active putative PAOs in the Westside Regional SIP incubations, and had

comparable AFEs to other EBPR-relevant members in the 'Penticton+propionate' incubation (Figure 2).

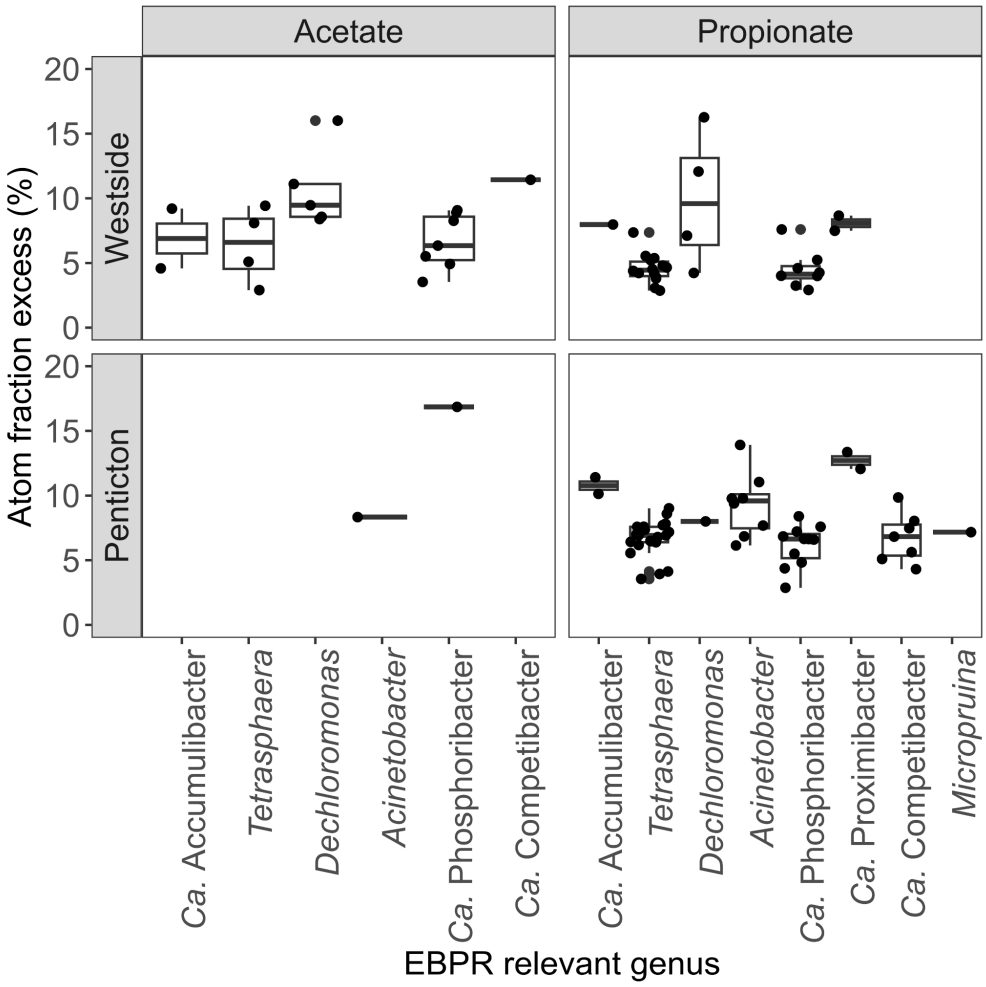


Figure 2. Atom fraction excess (AFE) of active EBPR-relevant 16S rRNA ASVs in the four SIP incubations. Boxplots indicate the distribution of AFEs for the ASVs within the same genera, and the points indicate the lower bounds of the FCR adjusted 95% confidence interval of bootstrapped (10,000 times) AFEs.

qSIP metagenomics on the 'Westside+acetate' SIP incubations

We recovered 385 de-replicated MAGs using a consensus binning approach with short-read metagenomic data from 95 individual buoyant density gradient fractions originating from ^{12}C and ^{13}C triplicate 'Westside+acetate' SIP incubations, along with long-read metagenome sequencing

of time-zero unlabelled Westside Regional bulk biomass. The mean overall alignment rate of the quality filtered shotgun metagenomic reads to the recovered de-replicated MAGs across the 95 buoyant density gradient fractions was 51.8 ± 4.8%, and was comparable between the ^{12}C (51.4 ± 1%) and ^{13}C (52.1 ± 0.5%) replicates (Datasheet S6). Out of these, 208 MAGs spanning 17 phyla (Figure S8) were retained for downstream analysis. These MAGs had estimated completeness >65% and contamination <12% (Datasheet S8). This threshold was operationally defined to retain the greatest number of EBPR-relevant MAGs for analysis. There were no noticeable differences in the relative abundances of the top 30 taxa between the ^{12}C and ^{13}C triplicates (Figure S9).

We estimated the absolute abundance of MAGs across buoyant density gradients using sequins as internal quantification standards[24] (Datasheet S11). No indications of anomalous sample processing or handling were observed, which was confirmed by the presence of the expected distribution of abundance profiles and peak buoyant densities for the pre-centrifugation spike-ins used as quality controls (Figure S10; Datasheet S1). Overall, 66 isotope-incorporating MAGs were identified across 15 phyla, with AFEs ranging from 1% to 11% (Figure 3; Datasheet S11). Several MAGs classified to the EBPR-relevant genera, *Ca*. Accumulibacter and *Ca.* Phosphoribacter, were metabolically active (Figure 3). These MAGs were also found to be relatively abundant across gradient fractions (>1% cumulative relative abundance) (Figure S11), similar to the observations from the 16S rRNA gene sequencing on the 'Westside+acetate' SIP incubations (Figure S6). The AFEs of active EBPR-relevant MAGs and 16S rRNA ASVs classified at the genus level were comparable (Figure S12).

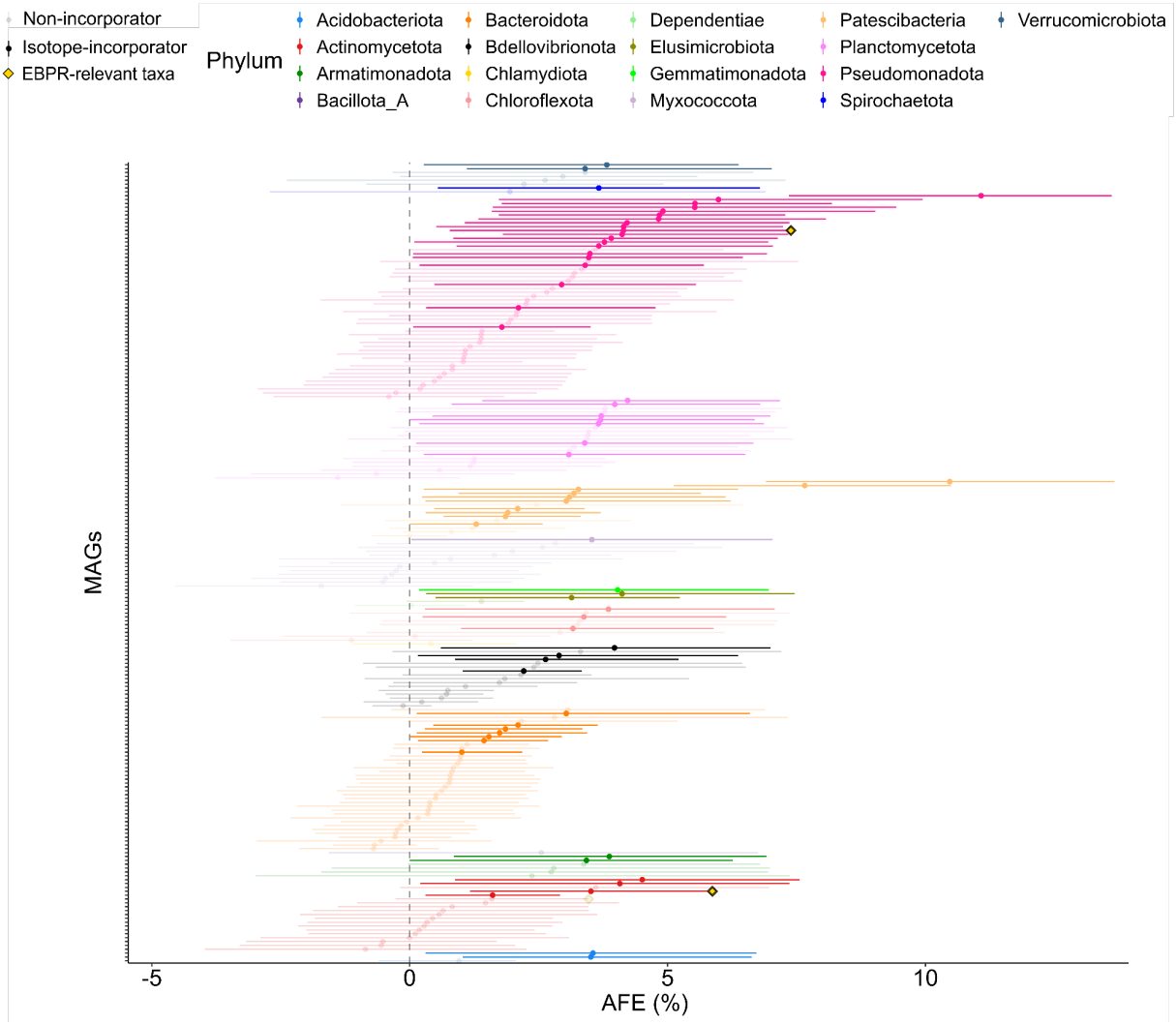


Figure 3. Isotopic labeling (i.e., atom fraction excess (AFE)) of 208 MAGs in the 'Westside+acetate' incubation with completeness >65% and contamination <12% across 17 phyla. EBPR-relevant taxa are identified with a gold diamond next to the error bars. Higher opacity indicates MAG status as metabolically active, i.e., the FCR-adjusted lower 95% confidence interval of AFE was >0. For details, see Datasheet S11. One non-incorporator MAG (MAG_qSIP_UBC.258) classified as *Citrobacter braaki* was removed from this plot for visualization purposes, due to its large range in the AFE in the bootstrapped data (lower 95% CI of -34%, and upper 95% CI of 6.8%).

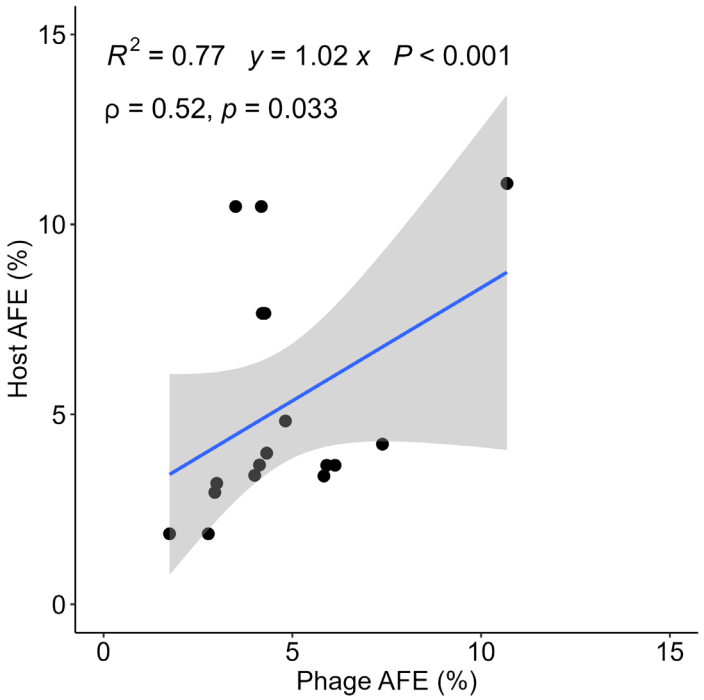
Despite the detection of ^{13}C-labelled *Dechloromonas* 16S rRNA ASVs (Figure 2), metagenomic binning did not recover any *Dechloromonas* genomes, which was likely due to their low abundance in the bulk biomass. We therefore mapped metagenomic reads from the SIP incubations to available reference *Dechloromonas* genomes to determine their presence. Two

known PAOs belonging to *Ca*. Dechloromonas phosphorivorans[59] (NCBI accessions: GCA_016714975 and GCA_016709495) had markedly higher coverage than other *Dechloromonas* reference genomes (Figure S13), indicating that they were likely present in the Westside Regional biomass. Yet, the mean coverage values for these two genomes were still below 1.0 (Figure S13), supporting the notion that low coverage was the reason a *Dechloromonas* MAG was not recovered. Further, a high degree of similarity was found between the metabolically active *Dechloromonas*-related 16S rRNA gene ASVs within the 'Westside+acetate' incubation and the 16S rRNA gene of *Ca*. Dechloromonas phosphorivorans, as evidenced by BLASTn analysis (99.7% identity with e-value of 0 for GCA_016714975 with ASV_591, and 97.6% identity with e-value of 0 for GCA_016709495 with ASV_3847), suggesting a likely presence of *Ca*. Dechloromonas phosphorivorans in the community. By mapping metagenomic reads from SIP buoyant density fractions to the two *Ca*. Dechloromonas phosphorivorans reference genomes, both GCA_016714975 and GCA_016709495 were determined to be metabolically active within the 'Westside+acetate' incubation, with AFEs of 6.3% and 4.3%, respectively. This observation of *Dechloromonas* taxa with comparatively higher AFE among known EBPR-relevant members was consistent with the 16S rRNA gene-based qSIP approach on the same 'Westside+acetate' batch (Figure 2). The related 16S rRNA gene ASVs of GCA_016714975 and GCA_016709495 (ASV_591 and ASV_3847, respectively) had measured AFEs of 8.6% and 9.7% in the 'Westside+acetate' batch test, respectively, using 16S rRNA gene-based qSIP.

Notably, the two aforementioned *Dechloromonas* genomes were predicted by iPHoP[47] as likely hosts for isotopically labelled phages recovered from the qSIP metagenomic data. Here, GCA_016709495 was identified as the host of an unclassified phage with a temperate lifestyle (AFE: 4.5%), while GCA_016714975 was identified as the host of another unclassified phage with an unknown lifestyle (AFE: 8.3%) (Datasheet S12). In addition to these two reference genomes, members belonging to the *Dechloromonas* genus were also predicted to be hosts of four other

390 isotopically labelled, unclassified phages (Datasheet S12). One of these phages was predicted
391 by PhaTYP[49] to have a temperate lifestyle, while the remaining three have unknown lifestyles.

392 Overall, 621 phage genomes were identified as metabolically active out of 5099 identified quality
393 filtered phages (Text S3). Out of these, 17 metabolically active phages were host-associated with
394 13 quality filtered metabolically active MAGs recovered in this study. There was a statistically
395 significant Spearman correlation (R=0.52; *p*-value<0.05), and a statistically significant positive
396 relationship (R^2=0.81; *p*-value<0.001) between the AFE of metabolically active MAGs and the
397 active phages they were predicted to host (Figure 4). That is, bacterial MAGs with higher
398 enrichment tended to have putative phage with higher enrichment. Several metabolically active
399 phages were associated with the genera of EBPR-relevant MAGs recovered in this study. Three
400 isotopically labelled phages were associated with the *Ca*. Accumulibacter genus (Datasheet S12).
401 Two of these three phages were predicted to have a temperate lifestyle, in which the phage inserts
402 its genome into the host indefinitely, with one classified as *Zierdtviridae* (AFE: 6%) and the other
403 yet-to-be-classified (AFE: 6.6%). The lifestyle of the third unclassified phage (AFE: 4.4%) could
404 not be predicted (Datasheet S12). Similarly, five metabolically active phages (AFE ranging
405 between 3.2% and 10.2%) were predicted that hit members within the *Ca.* Phosphoribacter genus
406 (Datasheet S12). One of these phages belonged to the *Vilmaviridae* family and was predicted to
407 have a temperate lifestyle, while the remaining four were unclassified with unknown lifestyles
408 (Datasheet S12). Also, an unclassified and unknown lifestyle isotopically labelled phage (AFE:
409 7.2%) was associated with *Ca*. Competibacter_A (Datasheet S12), a genus harboring known
410 GAOs[60]. However, no MAG classified as *Ca.* Competibacter_A was recovered in this study.



411

412 **Figure 4**: Comparison of AFEs of isotopically labelled phage and their predicted host MAGs.
413 Spearman correlation coefficient is indicated with ρ and *p-value* for Spearman correlation is
414 indicated with 'p'. Coefficient of determination is indicated with R^2 and *p*-value for the F-value of
415 linear regression equation between host and phage AFE is indicated as 'P'. Phage-host
416 predictions were performed using iPHoP.

417

418 ***Metabolic traits of putative EBPR members related to P cycling in***
419 ***'Westside+acetate' SIP incubations***

420 Members from both *Tetrasphaera*-associated groups (e.g., *Phycicoccus_A* and *Ca.*
421 Phosphoribacter) were present in the biomass of 'Westside+acetate' SIP incubations. However,
422 while *Phycicoccus_A_bin_qSIP_UBC.8* was abundant (sum of relative abundance across
423 fractions within ^{12}C and ^{13}C replicates: 1.1% ± 0.1% and 1.1% ± 0.02%), it did not anaerobically
424 uptake acetate (AFE: 0.1% and FCR adjusted lower 95% CI: -2%) (Datasheet S11). Indeed,
425 genome annotation revealed that *Phycicoccus_A_bin_qSIP_UBC.8* did not encode for acetate
426 permease (*actP*) and acetate-CoA ligase (*acs*), which play a key role in acetate metabolism for
427 growth via active transport of acetate intracellularly[61] and the conversion of acetate to acetyl-CoA
428 respectively[62]. Moreover, *Phycicoccus_A_bin_qSIP_UBC.8* shared an ANI of 98.3% with

Phycicoccus_A sp020161815 (GCA_020161815.1), which also did not encode for *actP*. However, *Phycicoccus_A_bin_qSIP_UBC.8* is likely an EBPR-relevant member by encoding for the low-affinity (*pit*) and high-affinity P transporters (*pstSABC*), P uptake regulator (*phoU*), polyphosphate kinase (*ppk*), and exopolyphosphatase (*ppx*) genes (Datasheet S13). Additionally, *Phycicoccus_A_bin_qSIP_UBC.8* was identified to encode glycogen biosynthesis genes (*glgB*, *glgC*, *glgP*, *glgY*, and *galU*) and did not encode for PHA synthase (*phaC*) (Datasheet S13), indicating potentially alternate sources of energy for growth.

Ca. Phosphoribacter members were observed to exhibit varied anaerobic acetate utilization in the 'Westside+acetate' incubations. For instance, *Phosphoribacter_bin_qSIP_UBC.2* was identified as an isotope incorporator (AFE: 3.5%). However, *actP*, *acs*, and *ackA* were not present and/or annotated within this MAG (Datasheet S13). The absence of these genes in *Phosphoribacter_bin_qSIP_UBC.2* may be attributed to its medium genome completion (65.02%) (Datasheet S8). *Phosphoribacter_bin_qSIP_UBC.3* encoded genes for acetate transport and utilization (Datasheet S13), but was not identified as an isotope incorporator (AFE: 1.6%; FCR adjusted lower 95% CI: -0.3%) (Datasheet S11). Similar to *Phycicoccus_A*, *Ca*. Phosphoribacter MAGs encoded for glycogen biosynthesis and P cycling genes, but did not encode *phaC* (Datasheet S13). Additionally, a high-completion (100%) and low-contamination (0.63%) isotopically labelled *Ca*. Accumulibacter MAG (*Accumulibacter_bin_qSIP_UBC.1*; Datasheet S8) had an AFE of 4.15% (Datasheet S11), and was observed to display all hallmark genes relevant to P and PHA cycling, glycogen biosynthesis, and encoded for the acetate transport and utilization genes (Datasheet S13).

To identify potentially novel or uncharacterized PAOs in the 'Westside+acetate' SIP incubations, isotopic incorporators were hierarchically clustered based on the presence or absence of genes thought to be relevant for P cycling in EBPR (e.g. inorganic P and PHA cycling, acetate uptake, glycogen biosynthesis, and fatty acid utilization); Datasheet S13). This analysis revealed clusters

454 resembling either an “Accumulibacter-type model” of P cycling (i.e., with both P and PHA cycling-
455 related genomic potential) or a “*Tetrasphaera*-type model” (i.e., containing only P cycling and
456 lacking PHA cycling-related genomic potential) (Figure 5).

457 Within the cluster of MAGs affiliated with the “Accumulibacter-type model” were two MAGs from
458 the family *Rhodocyclaceae*, including the *Ca*. Accumulibacter MAG
459 (*Accumulibacter_bin_qSIP_UBC.1*, AFE = 4%) and a MAG belonging to the genus JADJEF01
460 (*bin_qSIP_UBC.17*; AFE = 4%). Also included were several MAGs that were relatively highly
461 labelled from the order *Burkholderiales*, including a MAG (*bin_qSIP_UBC.84*; AFE = 6%) within
462 the genus JAABRC01 (NCBI classification = *Rhodoferax*, recently renamed as *Egaibacter*[63]), a
463 MAG classified to the genus CAIKVZ01 (*bin_qSIP_UBC.270*; AFE = 5%), and a MAG classified
464 to the genus *Casimicrobium* (*bin_qSIP_UBC.121*; AFE = 5%). Additionally, two MAGs classified
465 to the genus *Marmoricola* (phylum = *Actinomycetota*, family = *Nocardioidaceae*) had AFEs
466 between 4% and 5% and were also clustered within the “Accumulibacter-type model”.

467 The “*Tetrasphaera*-type model” cluster included the *Ca*. Phosphoribacter MAG
468 (*Phosphoribacter_bin_qSIP_UBC.2*; AFE=4%), a MAG assigned to the family
469 *Gemmatimonadaceae* (*bin_qSIP_UBC.97*, AFE = 4%), and two MAGs from the class
470 *Verrucomicrobiae* (*bin_qSIP_UBC.49* and *bin_qSIP_UBC.195*). Additionally, seven MAGs
471 associated with the phylum *Planctomycetes* fell within the “*Tetrasphaera*-type model” cluster, and
472 had AFEs between 3% to 4% (Figure 5).

473 A separate cluster of active MAGs was also observed that included members classified to
474 *Bdellovibrionota*, *Myxococcota*, *Acidobacteriota*, *Bacteroidota*, and *Actinomycetota*, which
475 generally contained several genes for P and PHA cycling (Figure 5). Therefore, this analysis
476 highlighted some acetate-assimilating MAGs that were similar to the known PAOs, *Tetrasphaera*

and *Ca.* Accumilibacter, based their P-cycling gene content, while it also uncovered some acetate assimilators with different metabolic repertoires.

Notably, the two of the five MAGs with the highest AFEs belonged to the order *Rickettsiales* (phylum = *Pseudomonadota*), which fell within a distinct cluster outside the "Accumulibacter-type model" and "*Tetrasphaera*-type model" types, suggesting they differed from these model PAOs in terms of P-cycling gene content (Figure 5). The MAG (*bin_qSIP_UBC.220*) with the highest observed AFE (11%; Figure 3; Datasheet S11) was classified to the QFOX01 genus within the UBA3002 family of *Rickettsiales* (Completeness: 100%, Contamination: 6.7%, Datasheet S8). This MAG contained many genes relevant for EBPR behavior, including acetate metabolism, P cycling, and PHA cycling (Figure 5). While it lacked *pit*, a low-affinity inorganic phosphate transporter, and *phoU*, a phosphate transport system regulatory protein, it did possess *pstSCAB* encoding for a high-affinity phosphate transporter (Figure 5). The other *Rickettsiales* MAG (*bin_qSIP_UBC.115*) within the top five highest incorporators (6% AFE; Datasheet S11) was also classified to the genus UBA3002 (Completeness: 93.4%, Contamination: 9.6%, Datasheet S8), and had a similar metabolic repertoire (Figure 5). Two distinct clusters of incorporators were also observed for MAGs belonging to the phylum *Patescibacteria* (Figure 5), within which the order *Saccharimonadales* harboured 3 of the top 5 highest incorporators (Figure 3; Datasheet S11). One highly labelled *Saccharimonadales* MAG (*bin_qSIP_UBC.283*; AFE = 8%) shared 97% ANI with the species *Ca.* Saccharimonas aalborgensis (Datasheet S11). Interestingly, the active MAGs within these *Patescibacteria* clusters did not contain many EBPR-relevant genes (Figure 5), suggesting an alternative mechanism for carbon labelling in these organisms.

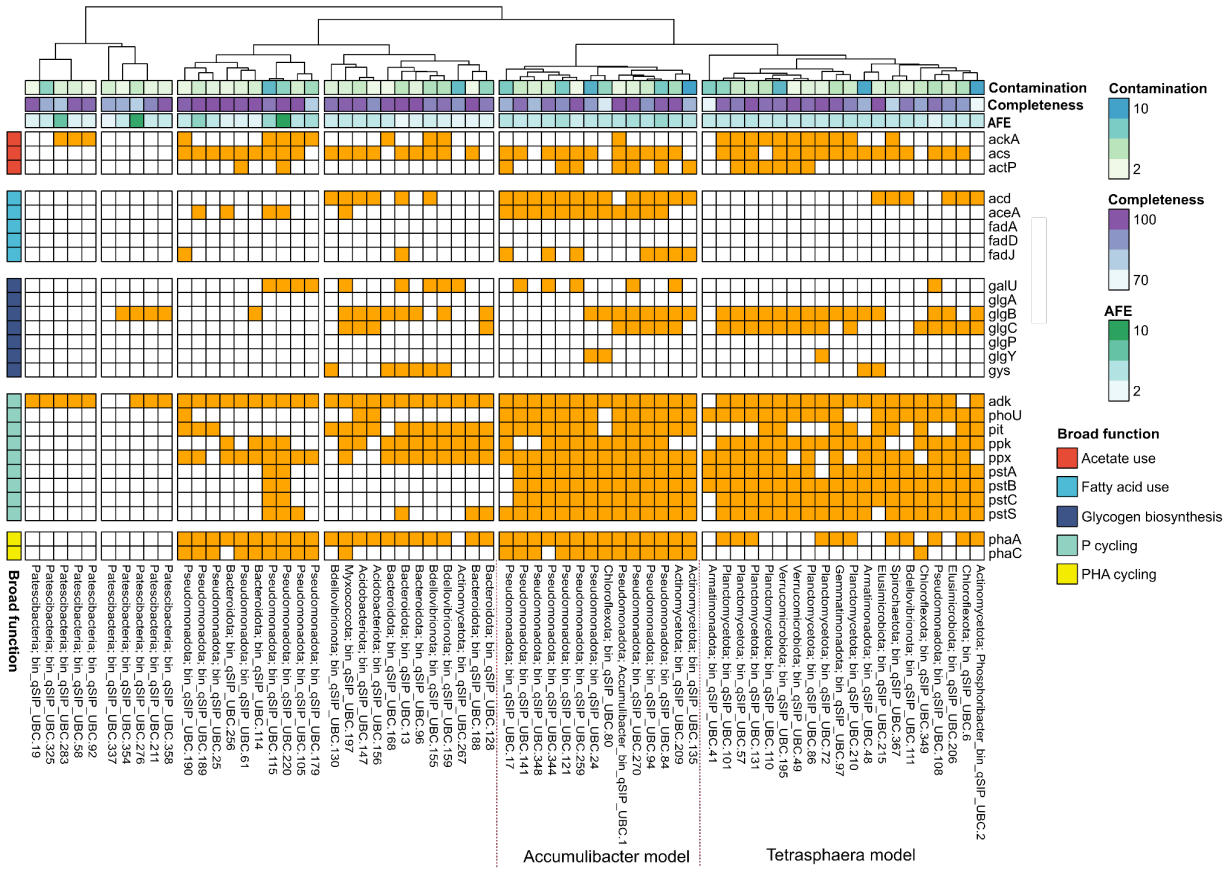


Figure 5. Hierarchical clustering of presence or absence of EBPR-relevant genes in the isotope incorporators. The unsupervised hierarchical clustering was performed with a "binary" distance metric and "ward.D2" linkage. The EBPR-relevant genes (listed on the right) are categorized based on broad function (shown on the left). The "*Ca*. Accumulibacter-type" and "*Tetrasphaera*-type" PAO models are identified at the bottom. AFE (%), completeness and contamination (%) (estimated using CheckM2) are shown as a heatmap on the top. Phylum names precede the MAG names with a ";". Gene abbreviations are as follows: *ackA* - acetate kinase; *acs* - acetyl-CoA synthetase; *actP* - acetate transporter; *acd* - acyl-CoA dehydrogenase / acetaldehyde dehydrogenase; *aceA* - isocitrate lyase; *fadA* - 3-ketoacyl-CoA thiolase; *fadD* - acyl-CoA synthetase; *fadJ* - multifunctional beta-oxidation protein (enoyl-CoA hydratase / 3-hydroxyacyl-CoA dehydrogenase); *galU* - UTP--glucose-1-phosphate uridylyltransferase; *glgA* - glycogen synthase; *glgB* - 1,4-alpha-glucan branching enzyme; *glgC* - glucose-1-phosphate adenylyltransferase; *glgP* - glycogen phosphorylase; *gys* - glycogen synthase; *adk* - adenylate kinase; *phoU* - phosphate transport system regulatory protein; *pit* - low-affinity inorganic phosphate transporter; *ppk* - polyphosphate kinase; *ppx* - exopolyphosphatase; *pstA* - high-affinity phosphate transport system permease protein PstA; *pstB* - high-affinity phosphate transport system ATP-binding protein PstB; *pstC* - high-affinity phosphate transport system permease protein PstC; *pstS* - high-affinity phosphate-binding protein PstS; *phaA* - acetyl-CoA acetyltransferase; *phaC* - poly(3-hydroxyalkanoate) polymerase

Discussion

This study investigated how anaerobic uptake of two VFAs, acetate and propionate, impacted microbial activity and P removal efficiency in biomass from two full-scale WRRFs. These two facilities use process configurations that promote internal VFA generation through primary sludge fermentation, a strategy specifically designed to enhance PAO growth and improve P removal efficiency[64]. Our findings indicate that the microbial communities within each facility were distinct, and they also differed in their anaerobic VFA uptake preference for efficient P removal during cyclic anaerobic/aerobic SIP incubations. DNA-SIP also revealed site-specific anaerobic carbon uptake preferences within known PAO taxonomic groups. For example, *Ca.* Accumulibacter and *Tetrasphaera*-related organisms in the Westside Regional biomass anaerobically utilized acetate as the carbon source for growth, while members of those groups in the Penticton biomass only consumed propionate. Our data suggest differences in earlier conflicting reports regarding VFA preference for efficient P cycling in EBPR systems, with some favoring acetate[7,8] and others propionate[9,10], were therefore likely driven by differences in microbial community composition and the environmental factors that shape them, such as immigration[65]. This highlights the utility of site-specific activity tests to determine the type of VFA that best promotes P uptake by the PAOs within a given WRRF, which could then be used to strategically manage sludge fermenters onsite to provide favorable growth substrates for efficient P cycling.

Even low abundance PAOs, such as *Dechloromonas*, exhibited high levels of metabolic activity and likely participated as key players driving P removal. A possible explanation for the low abundance of *Dechloromonas* members, despite their high activity in the ‘Westside+acetate’ incubation, could be attributed to active phage-induced lysis[20]. We identified a link between host metabolic activity and phage predation (indicated by isotopic labeling of phage genomes with qSIP metagenomics), potentially attributable to the "kill-the-winner" strategy involving phage-induced lysis[66]. Such a strategy would reduce the overall abundance of the host, thus obfuscating

their genome recovery within metagenomic data[67,68]. We observed that some phages infecting known PAOs had a predicted temperate lifestyle, wherein lysis could be triggered under specific environmental conditions[69]. Previous research suggests that environmental conditions, such as the presence of heavy metals or antibiotics, can induce prophages and lyse their PAO hosts, leading to decreased EBPR performance[70]. Barr et al.[71] attributed a crash in EBPR performance to prophage induced lysis and not to the competition with GAOs[71] that are historically thought to negatively impact P removal performance[72–74]. Overall, this study underscores the value of qSIP-metagenomics to identify host-phage linkages within active PAOs in wastewater treatment processes, which could inform further detailed experimentation to isolate and/or characterize such phages to elucidate their ecological roles within the EBPR processes.

Integrating qSIP with metagenomics also provides genomic evidence of substrate assimilation during a metabolism of interest[75], informing opportunities for further targeted physiological investigations of potentially novel organisms involved in a given function[76]. Here, utilizing a metabolic trait-based hierarchical clustering of isotopically labelled members in the 'Westside+acetate' anaerobic/aerobic SIP incubation, we identified active members encoding hallmark genes associated with P cycling in EBPR-related metabolism based on both "*Ca.* Accumulibacter-type"[77,78] and "*Tetrasphaera*-type" models[79,80]. The observation of an isotopically labelled *Rhodocyclaceae* MAG clustering with the metabolic traits of *Ca.* Accumulibacter aligns with previous observations of these members behaving as PAOs[81,82], and highlights the ambiguous boundary of the PAO phenotype within *Rhodocyclaceae*[83]*.* We also observed MAGs resembling similar traits as *Ca.* Accumulibacter that were related to *Rhodoferax*, which have been postulated as potential PAOs[63] and were observed to assimilate acetate[84], although aerobically. The isotopic labeling of members of *Gemmatimonadaceae* that had similar metabolic traits as the *Tetrasphaera*-type model of P cycling supports previous reports of this group harboring PAOs[85,86]. We also observed anaerobic carbon assimilation by multiple MAGs belonging to *Planctomycetes*,

569 which have not been previously reported as PAOs, yet were clustered with known PAOs within
570 the "*Tetrasphaera*-type" model for P cycling. While their role in EBPR processes is not clear,
571 *Planctomycetes* members have been associated with inorganic phosphorus solubilization in
572 soils[87], indicating a potential for their participation in P cycling.

573 Interestingly, the highest levels of anaerobic carbon assimilation were attributed to members of
574 *Saccharimonadales* (*Patescibacteria*) and *Rickettsiales* (*Pseudomonadota*), which formed
575 distinct clusters based on metabolic P cycling genes that were distant from both the "*Ca.*
576 Accumulibacter-type" and "*Tetrasphaera*-type" models. Members of these two groups have been
577 identified as host-associated, with several isolated members of *Saccharimonadales* displaying a
578 parasitic lifestyle with hosts from the phylum *Actinomycetota*[88–91], while *Rickettsiales* are often
579 endosymbiotic or parasitic with a variety of eukaryotes[92,93]. Higher abundances of
580 *Saccharimonadales* have been associated with better P removal within an EBPR system[94], and
581 a recent global survey of 16S rRNA genes in WRRFs identified a statistical correlation between
582 members of *Saccharimonadia* (the class in which *Saccharimonadales* falls) and *Tetrasphaera*-
583 related organisms, including *Ca.* Phosphoribacter[95]. Interestingly, the labelled
584 *Saccharimonadales* MAGs did not contain nearly any hallmark P cycling genes, and only a single
585 gene associated with acetate metabolism (acetate kinase) was observed in three active
586 *Saccharimonadales* MAGs. One highly labelled *Saccharimonadales* MAG was identified as the
587 species *Ca.* Saccharimonas aalborgensis, which has been proposed to have a limited metabolic
588 repertoire and an obligatory fermentative metabolism[96]. Collectively, this highlights the possibility
589 that *Saccharimonadales* MAGs could have assimilated carbon during the anaerobic/aerobic SIP
590 microcosms through a host-association with an active *Actinomycetota*, potentially the MAG
591 (Phosphoribacter_bin_qSIP_UBC.2) belonging to the *Ca*. Phosphoribacter genus that was
592 previously classified as *Tetrasphaera*[79]. In contrast to *Saccharimonadales*, the highly labelled
593 members of *Rickettsiales* possessed a variety of genes for P cycling and PHA storage, and could

therefore represent a yet to be characterized group of PAOs. While many *Rickettsiales* reproduce intracellularly with a eukaryotic host[93], the genus UBA3002 that was highly labelled in the anaerobic/aerobic SIP incubations is thought to represent an evolutionary 'basal' lineage that is free-living[97]. Nonetheless, it is also possible that UBA3002 MAGs were labeled through an association with a eukaryote that predated on active PAOs[98] or assimilated VFAs. Therefore, while carbon assimilation in these SIP microcosms does not confirm a PAO physiology, their labeling and metabolic repertoire warrants further targeted physiological assessments, for instance with Raman-FISH, to validate whether these active members store and cycle intracellular P[99].

Overall, the application of DNA qSIP with multiple VFAs under alternating anaerobic/aerobic conditions in this study highlights the complex microbial interactions in communities capable of cycling carbon and P under alternating redox conditions. We found that anaerobic VFA preference can differ within known groups of PAOs across different treatment plants, and the abundance of PAOs may be modulated by phages that actively infect these populations. In addition, there may be carbon cross-feeding between PAO hosts and bacterial parasites or predators, as is posited for the case of *Saccharimonadales* and *Ca*. Phosphoribacter. Finally, there is evidence that several groups of organisms that have so-far not been characterized as PAOs participate in anaerobic/aerobic carbon cycling and have metabolic potential to cycle P, highlighting the currently unexplored diversity of this important metabolic guild central to biological nutrient removal processes.

Acknowledgements

This work was performed under the auspices of the Natural Sciences and Engineering Research Council (NSERC) of Canada Discovery Grants program (RGPIN-2018-04585 to R.M.Z.) and the Genome British Columbia Sector Innovation Program (SIP06 to R.M.Z.). The work (Community Science Project ID: 505410 to R.M.Z.; https://doi.org/10.46936/10.25585/60001269) conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. We thank the staff at the Westside Regional and Penticton WRRFs, as well as Dr. Piero Galvagno, for kindly assisting with sample collection and delivery. We thank Dr. David Levy-Booth for his kind instruction and support in SIP fractionation.

Supporting Information

Supplementary Text S1 – S4: Additional methodological details on WRRF sampling, SIP incubations, DNA density gradient centrifugation, and bioinformatics.

Supplementary Figures S1-S13: WRRF configuration, bioinformatics workflow, microbial community abundance data, metagenomic data, SIP quality control data.

Supplementary Tables S1-S4: PCR protocol, SIP incubation nutrient data, isotopic labeling of bulk DNA.

Supplementary Datasheets S1-S14: NCBI accessions, MAG metadata, metagenomic coverage values across density gradients, gene annotation and hierarchal clustering, and details on SIP quality control spike ins.

References

(1) Albertsen, M.; Hansen, L. B. S.; Saunders, A. M.; Nielsen, P. H.; Nielsen, K. L. A Metagenome of a Full-Scale Microbial Community Carrying out Enhanced Biological Phosphorus Removal. *ISME J.* **2012**, *6* (6), 1094–1106. https://doi.org/10.1038/ismej.2011.176.

(2) Panswad, T.; Doungchai, A.; Anotai, J. Temperature Effect on Microbial Community of Enhanced Biological Phosphorus Removal System. *Water Res.* **2003**, *37* (2), 409–415. https://doi.org/10.1016/S0043-1354(02)00286-5.

(3) Randall, A. A.; Benefield, L. D.; Hill, W. E.; Nicol, J.-P.; Boman, G. K.; Jingo, S.-R. The Effect of Volatile Fatty Acids on Enhanced Biological Phosphorus Removal and Population Structure in Anaerobic/Aerobic Sequencing Batch Reactors. *Seq. Batch React. Technol. Batch Appl. Period. Unsteady-State Process.* **1997**, *35* (1), 153–160. https://doi.org/10.1016/S0273-1223(96)00891-8.

(4) Barnard, J. L.; Dunlap, P.; Steichen, M. Rethinking the Mechanisms of Biological Phosphorus Removal: Barnard et Al. *Water Environ. Res.* **2017**, *89* (11), 2043–2054.

(5) Gu, A.; Tooker, N.; Onnis-Hayden, A.; Wang, D.; Srinivasan, V.; Li, G.; Takács, I.; Vargas, E. Optimization and Design of a Side-Stream EBPR Process as a Sustainable Approach for Achieving Stable and Efficient Phosphorus Removal. *Water Res. Found. DC USA* **2019**.

(6) Wang, D.; Tooker, N. B.; Srinivasan, V.; Li, G.; Fernandez, L. A.; Schauer, P.; Menniti, A.; Maher, C.; Bott, C. B.; Dombrowski, P.; Barnard, J. L.; Onnis-Hayden, A.; Gu, A. Z. Side-Stream Enhanced Biological Phosphorus Removal (S2EBPR) Process Improves System Performance - A Full-Scale Comparative Study. *Water Res.* **2019**, *167*, 115109. https://doi.org/10.1016/j.watres.2019.115109.

(7) Cai, W.; Huang, W.; Li, H.; Sun, B.; Xiao, H.; Zhang, Z.; Lei, Z. Acetate Favors More Phosphorus Accumulation into Aerobic Granular Sludge than Propionate during the Treatment of Synthetic Fermentation Liquor. *Bioresour. Technol.* **2016**, *214*, 596–603. https://doi.org/10.1016/j.biortech.2016.05.015.

(8) Pijuan, M.; Saunders, A. m.; Guisasola, A.; Baeza, J. a.; Casas, C.; Blackall, L. l. Enhanced Biological Phosphorus Removal in a Sequencing Batch Reactor Using Propionate as the Sole Carbon Source. *Biotechnol. Bioeng.* **2004**, *85* (1), 56–67. https://doi.org/10.1002/bit.10813.

(9) Chen, Y.; Randall, A. A.; McCue, T. The Efficiency of Enhanced Biological Phosphorus Removal from Real Wastewater Affected by Different Ratios of Acetic to Propionic Acid. *Water Res.* **2004**, *38* (1), 27–36. https://doi.org/10.1016/j.watres.2003.08.025.

(10) Oehmen, A.; Yuan, Z.; Blackall, L. L.; Keller, J. Short-Term Effects of Carbon Source on the Competition of Polyphosphate Accumulating Organisms and Glycogen Accumulating Organisms. *Water Sci. Technol.* **2004**, *50* (10), 139–144. https://doi.org/10.2166/wst.2004.0629.

(11) Oehmen, A.; Zeng, R. J.; Yuan, Z.; Keller, J. Anaerobic Metabolism of Propionate by Polyphosphate-Accumulating Organisms in Enhanced Biological Phosphorus Removal Systems. *Biotechnol. Bioeng.* **2005**, *91* (1), 43–53. https://doi.org/10.1002/bit.20480.

(12) Nielsen, P. H.; Mielczarek, A. T.; Kragelund, C.; Nielsen, J. L.; Saunders, A. M.; Kong, Y.; Hansen, A. A.; Vollertsen, J. A Conceptual Ecosystem Model of Microbial Communities in Enhanced Biological Phosphorus Removal Plants. *Water Res.* **2010**, *44* (17), 5070–5088. https://doi.org/10.1016/j.watres.2010.07.036.

(13) Wang, L.; Shen, N.; Oehmen, A.; Zhou, Y. The Impact of Temperature on the Metabolism of Volatile Fatty Acids by Polyphosphate Accumulating Organisms (PAOs). *Environ. Res.*

2020, *188*, 109729. https://doi.org/10.1016/j.envres.2020.109729.
(14) Kong, Y.; Nielsen, J. L.; Nielsen, P. H. Identity and Ecophysiology of Uncultured Actinobacterial Polyphosphate-Accumulating Organisms in Full-Scale Enhanced Biological Phosphorus Removal Plants. *Appl. Environ. Microbiol.* **2005**, *71* (7), 4076–4085. https://doi.org/10.1128/AEM.71.7.4076-4085.2005.
(15) Guo, Y.; Zeng, W.; Li, N.; Peng, Y. Effect of Electron Acceptor on Community Structures of Denitrifying Polyphosphate Accumulating Organisms in Anaerobic-Anoxic-Oxic (A2O) Process Using DNA Based Stable-Isotope Probing (DNA-SIP). *Chem. Eng. J.* **2018**, *334*, 2039–2049. https://doi.org/10.1016/j.cej.2017.11.170.
(16) Waldrop, M. M. News Feature: Microbes for Better Sewage Treatment. *Proc. Natl. Acad. Sci.* **2021**, *118* (32). https://doi.org/10.1073/pnas.2112863118.
(17) Singleton, C. M.; Petriglieri, F.; Kristensen, J. M.; Kirkegaard, R. H.; Michaelsen, T. Y.; Andersen, M. H.; Kondrotaite, Z.; Karst, S. M.; Dueholm, M. S.; Nielsen, P. H.; Albertsen, M. Connecting Structure to Function with the Recovery of over 1000 High-Quality Metagenome-Assembled Genomes from Activated Sludge Using Long-Read Sequencing. *Nat. Commun.* **2021**, *12* (1), 2009. https://doi.org/10.1038/s41467-021-22203-2.
(18) Dumont, M. G.; Murrell, J. C. Stable Isotope Probing — Linking Microbial Identity to Function. *Nat. Rev. Microbiol.* **2005**, *3* (6), 499–504. https://doi.org/10.1038/nrmicro1162.
(19) Hungate, B. A.; Mau, R. L.; Schwartz, E.; Caporaso, J. G.; Dijkstra, P.; Gestel, N. van; Koch, B. J.; Liu, C. M.; McHugh, T. A.; Marks, J. C.; Morrissey, E. M.; Price, L. B. Quantitative Microbial Ecology through Stable Isotope Probing. *Appl. Environ. Microbiol.* **2015**, *81* (21), 7570–7581. https://doi.org/10.1128/AEM.02280-15.
(20) Greenlon, A.; Sieradzki, E.; Zablocki, O.; Koch, B. J.; Foley, M. M.; Kimbrel, J. A.; Hungate, B. A.; Blazewicz, S. J.; Nuccio, E. E.; Sun, C. L. Quantitative Stable-Isotope Probing (qSIP) with Metagenomics Links Microbial Physiology and Activity to Soil Moisture in Mediterranean-Climate Grassland Ecosystems. *Msystems* **2022**, *7* (6), e00417-22.
(21) Sieradzki, E. T.; Koch, B. J.; Greenlon, A.; Sachdeva, R.; Malmstrom, R. R.; Mau, R. L.; Blazewicz, S. J.; Firestone, M. K.; Hofmockel, K. S.; Schwartz, E. Measurement Error and Resolution in Quantitative Stable Isotope Probing: Implications for Experimental Design. *Msystems* **2020**, *5* (4), e00151-20.
(22) Vyshenska, D.; Sampara, P.; Singh, K.; Tomatsu, A.; Kauffman, W. B.; Nuccio, E. E.; Blazewicz, S. J.; Pett-Ridge, J.; Louie, K. B.; Varghese, N. A Standardized Quantitative Analysis Strategy for Stable Isotope Probing Metagenomics. *Msystems* **2023**, e01280-22.
(23) Nuccio, E. E.; Blazewicz, S. J.; Lafler, M.; Campbell, A. N.; Kakouridis, A.; Kimbrel, J. A.; Wollard, J.; Vyshenska, D.; Riley, R.; Tomatsu, A.; Hestrin, R.; Malmstrom, R. R.; Firestone, M.; Pett-Ridge, J. HT-SIP: A Semi-Automated Stable Isotope Probing Pipeline Identifies Cross-Kingdom Interactions in the Hyphosphere of Arbuscular Mycorrhizal Fungi. *Microbiome* **2022**, *10* (1), 199. https://doi.org/10.1186/s40168-022-01391-z.
(24) Hardwick, S. A.; Chen, W. Y.; Wong, T.; Kanakamedala, B. S.; Deveson, I. W.; Ongley, S. E.; Santini, N. S.; Marcellin, E.; Smith, M. A.; Nielsen, L. K.; Lovelock, C. E.; Neilan, B. A.; Mercer, T. R. Synthetic Microbe Communities Provide Internal Reference Standards for Metagenome Sequencing and Analysis. *Nat. Commun.* **2018**, *9* (1), 3096. https://doi.org/10.1038/s41467-018-05555-0.
(25) Neufeld, J. D.; Vohra, J.; Dumont, M. G.; Lueders, T.; Manefield, M.; Friedrich, M. W.; Murrell, J. C. DNA Stable-Isotope Probing. *Nat. Protoc.* **2007**, *2* (4), 860–866. https://doi.org/10.1038/nprot.2007.109.
(26) Parada, A. E.; Needham, D. M.; Fuhrman, J. A. Every Base Matters: Assessing Small Subunit rRNA Primers for Marine Microbiomes with Mock Communities, Time Series and Global Field Samples. *Environ. Microbiol.* **2016**, *18* (5), 1403–1414. https://doi.org/10.1111/1462-2920.13023.
(27) Simpson, A.; Wood-Charlson, E. M.; Smith, M.; Koch, B. J.; Beilsmith, K.; Kimbrel, J. A.;

Kellom, M.; Hunter, C. I.; Walls, R. L.; Schriml, L. M.; Wilhelm, R. C. MISIP: A Data Standard for the Reuse and Reproducibility of Any Stable Isotope Probing-Derived Nucleic Acid Sequence and Experiment. *GigaScience* **2024**, *13*, giae071. https://doi.org/10.1093/gigascience/giae071.
(28) Comeau, A. M.; Douglas, G. M.; Langille, M. G. Microbiome Helper: A Custom and Streamlined Workflow for Microbiome Research. *MSystems* **2017**, *2* (1), 10–1128.
(29) Apprill, A.; McNally, S.; Parsons, R.; Weber, L. Minor Revision to V4 Region SSU rRNA 806R Gene Primer Greatly Increases Detection of SAR11 Bacterioplankton. *Aquat. Microb. Ecol.* **2015**, *75*, 129–137.
(30) Cheng, H.; Concepcion, G. T.; Feng, X.; Zhang, H.; Li, H. Haplotype-Resolved de Novo Assembly Using Phased Assembly Graphs with Hifiasm. *Nat. Methods* **2021**, *18* (2), 170–175. https://doi.org/10.1038/s41592-020-01056-5.
(31) Cheng, H.; Jarvis, E. D.; Fedrigo, O.; Koepfli, K.-P.; Urban, L.; Gemmell, N. J.; Li, H. Haplotype-Resolved Assembly of Diploid Genomes without Parental Data. *Nat. Biotechnol.* **2022**, *40* (9), 1332–1335. https://doi.org/10.1038/s41587-022-01261-x.
(32) Li, H. New Strategies to Improve Minimap2 Alignment Accuracy. *Bioinformatics* **2021**, *37* (23), 4572–4574. https://doi.org/10.1093/bioinformatics/btab705.
(33) Kang, D. D.; Li, F.; Kirton, E.; Thomas, A.; Egan, R.; An, H.; Wang, Z. MetaBAT 2: An Adaptive Binning Algorithm for Robust and Efficient Genome Reconstruction from Metagenome Assemblies. *PeerJ* **2019**, *7*, e7359. https://doi.org/10.7717/peerj.7359.
(34) Wu, Y.-W.; Simmons, B. A.; Singer, S. W. MaxBin 2.0: An Automated Binning Algorithm to Recover Genomes from Multiple Metagenomic Datasets. *Bioinformatics* **2016**, *32* (4), 605–607. https://doi.org/10.1093/bioinformatics/btv638.
(35) Nissen, J. N.; Johansen, J.; Allesøe, R. L.; Sønderby, C. K.; Armenteros, J. J. A.; Grønbech, C. H.; Jensen, L. J.; Nielsen, H. B.; Petersen, T. N.; Winther, O.; Rasmussen, S. Improved Metagenome Binning and Assembly Using Deep Variational Autoencoders. *Nat. Biotechnol.* **2021**, *39* (5), 555–560. https://doi.org/10.1038/s41587-020-00777-4.
(36) Sieber, C. M. K.; Probst, A. J.; Sharrar, A.; Thomas, B. C.; Hess, M.; Tringe, S. G.; Banfield, J. F. Recovery of Genomes from Metagenomes via a Dereplication, Aggregation and Scoring Strategy. *Nat. Microbiol.* **2018**, *3* (7), 836–843. https://doi.org/10.1038/s41564-018-0171-1.
(37) Hofmeyr, S.; Egan, R.; Georganas, E.; Copeland, A. C.; Riley, R.; Clum, A.; Eloe-Fadrosh, E.; Roux, S.; Goltsman, E.; Buluç, A.; Rokhsar, D.; Oliker, L.; Yelick, K. Terabase-Scale Metagenome Coassembly with MetaHipMer. *Sci. Rep.* **2020**, *10* (1), 10689. https://doi.org/10.1038/s41598-020-67416-5.
(38) Langmead, B.; Salzberg, S. L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9* (4), 357–359. https://doi.org/10.1038/nmeth.1923.
(39) Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25* (16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
(40) Olm, M. R.; Brown, C. T.; Brooks, B.; Banfield, J. F. dRep: A Tool for Fast and Accurate Genomic Comparisons That Enables Improved Genome Recovery from Metagenomes through de-Replication. *ISME J.* **2017**, *11* (12), 2864–2868. https://doi.org/10.1038/ismej.2017.126.
(41) Chklovski, A.; Parks, D. H.; Woodcroft, B. J.; Tyson, G. W. CheckM2: A Rapid, Scalable and Accurate Tool for Assessing Microbial Genome Quality Using Machine Learning. *Nat. Methods* **2023**, *20* (8), 1203–1212. https://doi.org/10.1038/s41592-023-01940-w.
(42) Chaumeil, P.-A.; Mussig, A. J.; Hugenholtz, P.; Parks, D. H. GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database. *Bioinformatics* **2019**, btz848. https://doi.org/10.1093/bioinformatics/btz848.
(43) Parks, D. H.; Chuvochina, M.; Waite, D. W.; Rinke, C.; Skarshewski, A.; Chaumeil, P.-A.;

Hugenholtz, P. A Standardized Bacterial Taxonomy Based on Genome Phylogeny Substantially Revises the Tree of Life. *Nat. Biotechnol.* **2018**, *36* (10), 996–1004. https://doi.org/10.1038/nbt.4229.
(44) Shaffer, M.; Borton, M. A.; McGivern, B. B.; Zayed, A. A.; La Rosa, S. L.; Solden, L. M.; Liu, P.; Narrowe, A. B.; Rodríguez-Ramos, J.; Bolduc, B.; Gazitúa, M. C.; Daly, R. A.; Smith, G. J.; Vik, D. R.; Pope, P. B.; Sullivan, M. B.; Roux, S.; Wrighton, K. C. DRAM for Distilling Microbial Metabolism to Automate the Curation of Microbiome Function. *Nucleic Acids Res.* **2020**, *48* (16), 8883–8900. https://doi.org/10.1093/nar/gkaa621.
(45) Guo, J.; Bolduc, B.; Zayed, A. A.; Varsani, A.; Dominguez-Huerta, G.; Delmont, T. O.; Pratama, A. A.; Gazitúa, M. C.; Vik, D.; Sullivan, M. B.; Roux, S. VirSorter2: A Multi-Classifier, Expert-Guided Approach to Detect Diverse DNA and RNA Viruses. *Microbiome* **2021**, *9* (1), 37. https://doi.org/10.1186/s40168-020-00990-y.
(46) Nayfach, S.; Camargo, A. P.; Schulz, F.; Eloe-Fadrosh, E.; Roux, S.; Kyrpides, N. C. CheckV Assesses the Quality and Completeness of Metagenome-Assembled Viral Genomes. *Nat. Biotechnol.* **2021**, *39* (5), 578–585. https://doi.org/10.1038/s41587-020-00774-7.
(47) Roux, S.; Camargo, A. P.; Coutinho, F. H.; Dabdoub, S. M.; Dutilh, B. E.; Nayfach, S.; Tritt, A. iPHoP: An Integrated Machine Learning Framework to Maximize Host Prediction for Metagenome-Derived Viruses of Archaea and Bacteria. *PLoS Biol.* **2023**, *21* (4), e3002083.
(48) Shang, J.; Jiang, J.; Sun, Y. Bacteriophage Classification for Assembled Contigs Using Graph Convolutional Network. *Bioinformatics* **2021**, *37* (Supplement_1), i25–i33. https://doi.org/10.1093/bioinformatics/btab293.
(49) Shang, J.; Tang, X.; Sun, Y. PhaTYP: Predicting the Lifestyle for Bacteriophages Using BERT. *Brief. Bioinform.* **2023**, *24* (1), bbac487. https://doi.org/10.1093/bib/bbac487.
(50) Shang, J.; Peng, C.; Liao, H.; Tang, X.; Sun, Y. PhaBOX: A Web Server for Identifying and Characterizing Phage Contigs in Metagenomic Data. arXiv July 27, 2023. https://doi.org/10.48550/arXiv.2303.15707.
(51) Martin, M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet.journal* **2011**, *17* (1), 10–12. https://doi.org/10.14806/ej.17.1.200.
(52) Callahan, B. J.; McMurdie, P. J.; Rosen, M. J.; Han, A. W.; Johnson, A. J. A.; Holmes, S. P. DADA2: High-Resolution Sample Inference from Illumina Amplicon Data. *Nat. Methods* **2016**, *13* (7), 581–583. https://doi.org/10.1038/nmeth.3869.
(53) Morten Kam Dahl Dueholm; Kasper Skytte Andersen; Anne-Kirstine C. Petersen; Vibeke Rudkjøbing; MiDAS Global Consortium for Anaerobic Digesters; Per Halkjær Nielsen. MiDAS 5: Global Diversity of Bacteria and Archaea in Anaerobic Digesters. *bioRxiv* **2023**, 2023.08.24.554448. https://doi.org/10.1101/2023.08.24.554448.
(54) Dueholm, M. K. D.; Nierychlo, M.; Andersen, K. S.; Rudkjøbing, V.; Knutsson, S.; Albertsen, M.; Nielsen, P. H. MiDAS 4: A Global Catalogue of Full-Length 16S rRNA Gene Sequences and Taxonomy for Studies of Bacterial Communities in Wastewater Treatment Plants. *Nat. Commun.* **2022**, *13* (1), 1908. https://doi.org/10.1038/s41467-022-29438-7.
(55) Youngblut, N. D.; Barnett, S. E.; Buckley, D. H. HTSSIP: An R Package for Analysis of High Throughput Sequencing Data from Nucleic Acid Stable Isotope Probing (SIP) Experiments. *PLOS ONE* **2018**, *13* (1), e0189616. https://doi.org/10.1371/journal.pone.0189616.
(56) Venables, W. N.; Ripley, B. D. *Modern Applied Statistics with S*, Fourth.; Springer: New York, 2002.
(57) Vyshenska, D.; Sampara, P.; Singh, K.; Tomatsu, A.; Kauffman, W. B.; Nuccio, E. E.; Blazewicz, S. J.; Pett-Ridge, J.; Louie, K. B.; Varghese, N. A Standardized Quantitative Analysis Strategy for Stable Isotope Probing Metagenomics. *Msystems* **2023**, e01280-22.
(58) Hrenovic, J.; Tibljas, D.; Ivankovic, T.; Kovacevic, D.; Sekovanic, L. Sepiolite as Carrier of

the Phosphate-Accumulating Bacteria Acinetobacter Junii. *Appl. Clay Sci.* **2010**, *50* (4), 582–587. https://doi.org/10.1016/j.clay.2010.10.024.
(59) Petriglieri, F.; Singleton, C.; Peces, M.; Petersen, J. F.; Nierychlo, M.; Nielsen, P. H. "Candidatus Dechloromonas Phosphoritropha" and "Ca. D. Phosphorivorans", Novel Polyphosphate Accumulating Organisms Abundant in Wastewater Treatment Systems. *ISME J.* **2021**, *15* (12), 3605–3614. https://doi.org/10.1038/s41396-021-01029-2.
(60) McIlroy, S. J.; Albertsen, M.; Andresen, E. K.; Saunders, A. M.; Kristiansen, R.; Stokholm-Bjerregaard, M.; Nielsen, K. L.; Nielsen, P. H. 'Candidatus Competibacter'-Lineage Genomes Retrieved from Metagenomes Reveal Functional Metabolic Diversity. *ISME J.* **2014**, *8* (3), 613–624.
(61) Gimenez, R.; Nuñez, M. F.; Badia, J.; Aguilar, J.; Baldoma, L. The Gene yjcG, Cotranscribed with the Gene Acs, Encodes an Acetate Permease in Escherichia Coli. *J. Bacteriol.* **2003**, *185* (21), 6448–6455. https://doi.org/10.1128/jb.185.21.6448-6455.2003.
(62) Wolfe, A. J. The Acetate Switch. *Microbiol. Mol. Biol. Rev.* **2005**, *69* (1), 12–50. https://doi.org/10.1128/mmbr.69.1.12-50.2005.
(63) Petersen, J. F.; Valk, L. C.; Verhoeven, M. D.; Nierychlo, M. A.; Singleton, C. M.; Dueholm, M. K.; Nielsen, P. H. Diversity and Physiology of Abundant Rhodoferax Species in Global Wastewater Treatment Systems. *Syst. Appl. Microbiol.* **2025**, *48* (1), 126574.
(64) Chanona, J.; Ribes, J.; Seco, A.; Ferrer, J. Optimum Design and Operation of Primary Sludge Fermentation Schemes for Volatile Fatty Acids Production. *Water Res.* **2006**, *40* (1), 53–60.
(65) Vestergaard, S. Z.; Dottorini, G.; Peces, M.; Murguz, A.; Dueholm, M. K. D.; Nierychlo, M.; Nielsen, P. H. Microbial Core Communities in Activated Sludge Plants Are Strongly Affected by Immigration and Geography. *Environ. Microbiome* **2024**, *19* (1), 63. https://doi.org/10.1186/s40793-024-00604-2.
(66) Kirsch, J. M.; Brzozowski, R. S.; Faith, D.; Round, J. L.; Secor, P. R.; Duerkop, B. A. Bacteriophage-Bacteria Interactions in the Gut: From Invertebrates to Mammals. *Annu. Rev. Virol.* **2021**, *8* (1), 95–113. https://doi.org/10.1146/annurev-virology-091919-101238.
(67) Zhang, L.; Fang, X.; Liao, H.; Zhang, Z.; Zhou, X.; Han, L.; Chen, Y.; Qiu, Q.; Li, S. C. A Comprehensive Investigation of Metagenome Assembly by Linked-Read Sequencing. *Microbiome* **2020**, *8* (1), 156. https://doi.org/10.1186/s40168-020-00929-3.
(68) Shapiro, O. H.; Kushmaro, A.; Brenner, A. Bacteriophage Predation Regulates Microbial Abundance and Diversity in a Full-Scale Bioreactor Treating Industrial Wastewater. *ISME J.* **2010**, *4* (3), 327–336. https://doi.org/10.1038/ismej.2009.118.
(69) Howard-Varona, C.; Hargreaves, K. R.; Abedon, S. T.; Sullivan, M. B. Lysogeny in Nature: Mechanisms, Impact and Ecology of Temperate Phages. *ISME J.* **2017**, *11* (7), 1511–1520. https://doi.org/10.1038/ismej.2017.16.
(70) Motlagh, A. M.; Bhattacharjee, A. S.; Goel, R. Microbiological Study of Bacteriophage Induction in the Presence of Chemical Stress Factors in Enhanced Biological Phosphorus Removal (EBPR). *Water Res.* **2015**, *81*, 1–14. https://doi.org/10.1016/j.watres.2015.04.023.
(71) Barr, J. J.; Slater, F. R.; Fukushima, T.; Bond, P. L. Evidence for Bacteriophage Activity Causing Community and Performance Changes in a Phosphorus-Removal Activated Sludge. *FEMS Microbiol. Ecol.* **2010**, *74* (3), 631–642. https://doi.org/10.1111/j.1574-6941.2010.00967.x.
(72) Saunders, A. M.; Oehmen, A.; Blackall, L. L.; Yuan, Z.; Keller, J. The Effect of GAOs (Glycogen Accumulating Organisms) on Anaerobic Carbon Requirements in Full-Scale Australian EBPR (Enhanced Biological Phosphorus Removal) Plants. *Water Sci. Technol.* **2003**, *47* (11), 37–43. https://doi.org/10.2166/wst.2003.0584.
(73) Satoh, H.; Mino, T.; Matsuo, T. Deterioration of Enhanced Biological Phosphorus Removal by the Domination of Microorganisms without Polyphosphate Accumulation. *Water Sci.*

Technol. **1994**, *30* (6), 203–211. https://doi.org/10.2166/wst.1994.0270.
(74) Cech, J. S.; Hartman, P. Competition between Polyphosphate and Polysaccharide Accumulating Bacteria in Enhanced Biological Phosphate Removal Systems. *Water Res.* **1993**, *27* (7), 1219–1225. https://doi.org/10.1016/0043-1354(93)90014-9.
(75) Walkup, J.; Dang, C.; Mau, R. L.; Hayer, M.; Schwartz, E.; Stone, B. W.; Hofmockel, K. S.; Koch, B. J.; Purcell, A. M.; Pett-Ridge, J.; Wang, C.; Hungate, B. A.; Morrissey, E. M. The Predictive Power of Phylogeny on Growth Rates in Soil Bacterial Communities. *ISME Commun.* **2023**, *3* (1), 73. https://doi.org/10.1038/s43705-023-00281-1.
(76) Hatzenpichler, R.; Krukenberg, V.; Spietz, R. L.; Jay, Z. J. Next-Generation Physiology Approaches to Study Microbiome Function at Single Cell Level. *Nat. Rev. Microbiol.* **2020**, *18* (4), 241–256. https://doi.org/10.1038/s41579-020-0323-1.
(77) He, S.; McMahon, K. D. Microbiology of 'Candidatus Accumulibacter' in Activated Sludge. *Microb. Biotechnol.* **2011**, *4* (5), 603–619. https://doi.org/10.1111/j.1751-7915.2011.00248.x.
(78) Páez-Watson, T.; van Loosdrecht, M. C. M.; Wahl, S. A. From Metagenomes to Metabolism: Systematically Assessing the Metabolic Flux Feasibilities for "Candidatus Accumulibacter" Species during Anaerobic Substrate Uptake. *Water Res.* **2024**, *250*, 121028. https://doi.org/10.1016/j.watres.2023.121028.
(79) Singleton, C. M.; Petriglieri, F.; Wasmund, K.; Nierychlo, M.; Kondrotaite, Z.; Petersen, J. F.; Peces, M.; Dueholm, M. S.; Wagner, M.; Nielsen, P. H. The Novel Genus,'Candidatus Phosphoribacter', Previously Identified as Tetrasphaera, Is the Dominant Polyphosphate Accumulating Lineage in EBPR Wastewater Treatment Plants Worldwide. *ISME J.* **2022**, *16* (6), 1605–1616.
(80) Kristiansen, R.; Nguyen, H. T. T.; Saunders, A. M.; Nielsen, J. L.; Wimmer, R.; Le, V. Q.; McIlroy, S. J.; Petrovski, S.; Seviour, R. J.; Calteau, A.; Nielsen, K. L.; Nielsen, P. H. A Metabolic Model for Members of the Genus Tetrasphaera Involved in Enhanced Biological Phosphorus Removal. *ISME J.* **2013**, *7* (3), 543–554. https://doi.org/10.1038/ismej.2012.136.
(81) Klein, E.; Weiler, J.; Wagner, M.; Čelikić, M.; Niemeyer, C. M.; Horn, H.; Gescher, J. Enrichment of Phosphate-Accumulating Organisms (PAOs) in a Microfluidic Model Biofilm System by Mimicking a Typical Aerobic Granular Sludge Feast/Famine Regime. *Appl. Microbiol. Biotechnol.* **2022**, *106* (3), 1313–1324. https://doi.org/10.1007/s00253-022-11759-8.
(82) Zilles, J. L.; Peccia, J.; Kim, M.-W.; Hung, C.-H.; Noguera, D. R. Involvement of Rhodocyclus-Related Organisms in Phosphorus Removal in Full-Scale Wastewater Treatment Plants. *Appl. Environ. Microbiol.* **2002**, *68* (6), 2763–2769. https://doi.org/10.1128/AEM.68.6.2763-2769.2002.
(83) Oyserman, B. O.; Moya, F.; Lawson, C. E.; Garcia, A. L.; Vogt, M.; Heffernen, M.; Noguera, D. R.; McMahon, K. D. Ancestral Genome Reconstruction Identifies the Evolutionary Basis for Trait Acquisition in Polyphosphate Accumulating Bacteria. *ISME J.* **2016**, *10* (12), 2931–2945. https://doi.org/10.1038/ismej.2016.67.
(84) McIlroy, S. J.; Starnawska, A.; Starnawski, P.; Saunders, A. M.; Nierychlo, M.; Nielsen, P. H.; Nielsen, J. L. Identification of Active Denitrifiers in Full-Scale Nutrient Removal Wastewater Treatment Systems. *Environ. Microbiol.* **2016**, *18* (1), 50–64. https://doi.org/10.1111/1462-2920.12614.
(85) Hanada, S.; Sekiguchi, Y. The Phylum Gemmatimonadetes. In *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*; Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2014; pp 677–681. https://doi.org/10.1007/978-3-642-38954-2_164.
(86) Han, Z.; He, W.; Shi, H.; Wang, C.; Liu, C.; Feng, Y. Effects of the Integrated Use of Dairy Cow Manure on Soil Properties and Biological Fertility. *Sustainability* **2023**, *15* (15).

https://doi.org/10.3390/su151511693.
(87) Yuan, T.; Chen, S.; Zhang, Y.; Ji, L.; Dari, B.; Sihi, D.; Xu, D.; Zhang, Z.; Yan, Z.; Wang, X. Mechanism of Increased Soil Phosphorus Availability in a Calcareous Soil by Ammonium Polyphosphate. *Biol. Fertil. Soils* **2022**, *58* (6), 649–665. https://doi.org/10.1007/s00374-022-01650-z.
(88) Batinovic, S.; Rose, J. J. A.; Ratcliffe, J.; Seviour, R. J.; Petrovski, S. Cocultivation of an Ultrasmall Environmental Parasitic Bacterium with Lytic Ability against Bacteria Associated with Wastewater Foams. *Nat. Microbiol.* **2021**, *6* (6), 703–711. https://doi.org/10.1038/s41564-021-00892-1.
(89) Xie, B.; Wang, J.; Nie, Y.; Tian, J.; Wang, Z.; Chen, D.; Hu, B.; Wu, X.-L.; Du, W. Type IV Pili Trigger Episymbiotic Association of Saccharibacteria with Its Bacterial Host. *Proc. Natl. Acad. Sci.* **2022**, *119* (49), e2215990119. https://doi.org/10.1073/pnas.2215990119.
(90) He, X.; McLean, J. S.; Edlund, A.; Yooseph, S.; Hall, A. P.; Liu, S.-Y.; Dorrestein, P. C.; Esquenazi, E.; Hunter, R. C.; Cheng, G.; Nelson, K. E.; Lux, R.; Shi, W. Cultivation of a Human-Associated TM7 Phylotype Reveals a Reduced Genome and Epibiotic Parasitic Lifestyle. *Proc. Natl. Acad. Sci.* **2015**, *112* (1), 244–249. https://doi.org/10.1073/pnas.1419038112.
(91) Wang, Y.; Gallagher, L. A.; Andrade, P. A.; Liu, A.; Humphreys, I. R.; Turkarslan, S.; Cutler, K. J.; Arrieta-Ortiz, M. L.; Li, Y.; Radey, M. C.; McLean, J. S.; Cong, Q.; Baker, D.; Baliga, N. S.; Peterson, S. B.; Mougous, J. D. Genetic Manipulation of Patescibacteria Provides Mechanistic Insights into Microbial Dark Matter and the Epibiotic Lifestyle. *Cell* **2023**, *186* (22), 4803-4817.e13. https://doi.org/10.1016/j.cell.2023.08.017.
(92) Schön, M. E.; Martijn, J.; Vosseberg, J.; Köstlbacher, S.; Ettema, T. J. The Evolutionary Origin of Host Association in the Rickettsiales. *Nat. Microbiol.* **2022**, *7* (8), 1189–1199.
(93) Salje, J. Cells within Cells: Rickettsiales and the Obligate Intracellular Bacterial Lifestyle. *Nat. Rev. Microbiol.* **2021**, *19* (6), 375–390. https://doi.org/10.1038/s41579-020-00507-2.
(94) Zhang, J.; Lei, S.; Zhang, X.; Xie, S.; Zheng, Y.; Yang, W.; Wang, Z.; Chen, A.; Zhao, J. Enhanced Nitrogen and Phosphorus Removal by *Saccharimonadales* Sp. in a Sequencing Batch Reactor. *Biochem. Eng. J.* **2024**, *211*, 109456. https://doi.org/10.1016/j.bej.2024.109456.
(95) Hu, H.; Kristensen, J. M.; Herbold, C. W.; Pjevac, P.; Kitzinger, K.; Hausmann, B.; Dueholm, M. K. D.; Nielsen, P. H.; Wagner, M. Global Abundance Patterns, Diversity, and Ecology of Patescibacteria in Wastewater Treatment Plants. *Microbiome* **2024**, *12* (1), 55. https://doi.org/10.1186/s40168-024-01769-1.
(96) Albertsen, M.; Hugenholtz, P.; Skarshewski, A.; Nielsen, K. L.; Tyson, G. W.; Nielsen, P. H. Genome Sequences of Rare, Uncultured Bacteria Obtained by Differential Coverage Binning of Multiple Metagenomes. *Nat. Biotechnol.* **2013**, *31* (6), 533–538. https://doi.org/10.1038/nbt.2579.
(97) Moncadas, L. S.; Shabarova, T.; Kavagutti, V. S.; Bulzu, P.-A.; Chiriac, M.-C.; Park, S.-J.; Mukherjee, I.; Ghai, R.; Andrei, A.-S. Rickettsiales' Deep Evolutionary History Sheds Light on the Emergence of Intracellular Lifestyles. bioRxiv January 31, 2023, p 2023.01.31.526412. https://doi.org/10.1101/2023.01.31.526412.
(98) Čech, J. S.; Hartman, P.; Macek, M. Bacteria and Protozoa Population Dynamics in Biological Phosphate Removal Systems. *Water Sci. Technol.* **1994**, *29* (7), 109–117. https://doi.org/10.2166/wst.1994.0320.
(99) Fernando, E. Y.; McIlroy, S. J.; Nierychlo, M.; Herbst, F.-A.; Petriglieri, F.; Schmid, M. C.; Wagner, M.; Nielsen, J. L.; Nielsen, P. H. Resolving the Individual Contribution of Key Microbial Populations to Enhanced Biological Phosphorus Removal with Raman–FISH. *ISME J.* **2019**, *13* (8), 1933–1946. https://doi.org/10.1038/s41396-019-0399-7.