



Genome-Resolved Metagenomics of a Photosynthetic Bioreactor Performing Biological Nutrient Removal

 Elizabeth A. McDaniel,^a Renske Wever,^{c,d} Ben O. Oyserman,^{c,d} Daniel R. Noguera,^{b,e} Katherine D. McMahon^{a,b}

^aDepartment of Bacteriology, University of Wisconsin—Madison, Madison, Wisconsin, USA

^bDepartment of Civil and Environmental Engineering, University of Wisconsin—Madison, Madison, Wisconsin, USA

^cDepartment of Microbial Ecology, Netherlands Institute of Ecological Research, Wageningen, The Netherlands

^dBioinformatics Group, Wageningen University and Research, Wageningen, The Netherlands

^eDOE Great Lakes Bioenergy Research Center, Madison, Wisconsin, USA

ABSTRACT Enhanced biological phosphorus removal (EBPR) is an economically and environmentally significant wastewater treatment process for removing excess phosphorus by harnessing the metabolic physiologies of enriched microbial communities. We present a genome-resolved metagenomic data set consisting of 86 metagenome-assembled genome sequences from a photosynthetically operated lab-scale bioreactor simulating EBPR.

A photobioreactor performing light-supported enhanced biological phosphorus removal (EBPR) was seeded from two lab-scale wastewater treatment systems, a photosynthetic nitrifying enrichment culture and an activated sludge EBPR bioreactor designed to enrich for the polyphosphate-accumulating organism “*Candidatus Accumulibacter phosphatis*” (1, 2). The photobioreactor was designed to achieve phosphorus removal and denitrification without the addition of mechanical aeration by providing light (2). Three samples were collected during cycles 87, 103, and 129. DNA was extracted as described (2), and 100 ng of DNA was sheared to 300 bp using the Covaris LE220 ultrasonicator and size selected with solid-phase reversible immobilization (SPRI) beads (Beckman Coulter). The fragments were ligated with end repair, A-tailing, and Illumina-compatible adapters (IDT, Inc.) using the KAPA-Illumina library preparation kit (Kapa Biosystems). The libraries were quantified using the Kapa Biosystems next-generation sequencing library quantitative PCR (qPCR) kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified libraries were prepared using the v4 TruSeq paired-end cluster kit and the Illumina cBot instrument to create a clustered flow cell for sequencing. Shotgun metagenomic sequencing was performed on the Illumina HiSeq 2500 platform with the TruSeq sequencing-by-synthesis (SBS) sequencing kit, followed by 2 × 150-bp indexing. All metagenomic libraries consist of approximately 50 million 150-bp paired-end reads (table of metagenome information available at <https://doi.org/10.6084/m9.figshare.14164307>).

Default parameters were used for all software unless otherwise specified. The raw metagenomic reads were quality filtered and trimmed using bbdutk as part of the BBTools v38.07 suite (3). All three metagenomic samples were coassembled and individually assembled into contigs using SPAdes v3.9.0 with the metagenomic option (4). Metagenomic reads from each sample were mapped against all assemblies using BMap as part of the BBTools v38.07 suite with a 95% sequence identity cutoff (3). The assembled contigs from each assembly were binned into population genomes using MetaBAT2 v2.12.1 using only contigs larger than 1,000 bp (5). The bins were dereplicated using dRep v2.4.2 (6) to obtain the highest-quality representative set of genomes. This resulted in a total of 86 nonredundant species-resolved genome sequences, described in Table 1 and summarized in Fig. 1. All genome

Citation McDaniel EA, Wever R, Oyserman BO, Noguera DR, McMahon KD. 2021. Genome-resolved metagenomics of a photosynthetic bioreactor performing biological nutrient removal. Microbiol Resour Announc 10: e00244-21. <https://doi.org/10.1128/MRA.00244-21>.

Editor J. Cameron Thrash, University of Southern California

Copyright © 2021 McDaniel et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Elizabeth A. McDaniel, elizabethmcd93@gmail.com.

Received 12 March 2021

Accepted 31 March 2021

Published 6 May 2021

TABLE 1 Genome accession numbers and statistics^a

Relative abundance on date									
(mo/day/yr)									
Isolate name	Assembly accession no.	Code	Genome size (mbp)	Avg contig size (bp)	% GC	No. of contigs	No. of rRNAs	No. of tRNAs	BioProject accession no.
UWPOB_ACID01	GCA_017304135.1	ACID01	6.49	17,791	64.5	365	1	1	SAMN18059832 JAFB8000000000 PRJNA704939
UWPOB_ACID02	GCA_017303955.1	ACID02	6.15	133,664	34.5	46	0	0	SAMN18059833 JAFB8000000000 PRJNA704939
UWPOB_ARM01	GCA_017303935.1	ARMA1	2.56	511,480	53.2	5	1	1	SAMN18059834 JAFB8000000000 PRJNA704939
UWPOB_BACT1	GCA_017303995.1	BACT1	4.31	31,460	41.5	137	2	0	SAMN18059835 JAFB8000000000 PRJNA704939
UWPOB_BACT2	GCA_017303975.1	BACT2	3.47	46,223	35.2	75	3	1	SAMN18059837 JAFB8000000000 PRJNA704939
UWPOB_BACT3	GCA_017303895.1	BACT3	3.61	103,001	37.3	35	3	1	SAMN18059836 JAFB8000000000 PRJNA704939
UWPOB_CHIT1	GCA_017303905.1	CHIT1	3.95	564,856	36.8	7	2	1	SAMN18059838 JAFB8000000000 PRJNA704939
UWPOB_CHIT2	GCA_017303835.1	CHIT2	3.26	1,631,117	42.5	2	1	1	SAMN18059839 JAFB8000000000 PRJNA704939
UWPOB_CHIT3	GCA_017303415.1	CHIT3	3.72	1,239,140	46	3	0	1	SAMN18059840 JAFB8000000000 PRJNA704939
UWPOB_CHIT4	GCA_017303335.1	CHIT4	4.01	572,270	49.5	7	1	1	SAMN18059841 JAFB8000000000 PRJNA704939
UWPOB_CHIT5	GCA_017303455.1	CHIT5	4.55	79,888	43	57	0	0	SAMN18059843 JAFB8000000000 PRJNA704939
UWPOB_CHIT6	GCA_017303855.1	CHIT6	4.62	149,027	46.1	31	1	0	SAMN18059842 JAFB8000000000 PRJNA704939
UWPOB_CHIT7	GCA_017303395.1	CHIT7	3.76	83,566	41.9	45	0	0	SAMN18059844 JAFB8000000000 PRJNA704939
UWPOB_CHIT8	GCA_017303475.1	CHIT8	6.04	114,008	47.9	53	1	1	SAMN18059845 JAFB8000000000 PRJNA704939
UWPOB_CHIT9	GCA_017303875.1	CHIT9	5.78	90,342	50.4	64	0	0	SAMN18059846 JAFB8000000000 PRJNA704939
UWPOB_CYT01	GCA_017303805.1	CYT01	4.29	329,632	45.1	13	1	0	SAMN18059847 JAFB8000000000 PRJNA704939
UWPOB_CYT02	GCA_017303795.1	CYT02	4.87	180,295	42.2	27	1	0	SAMN18059848 JAFB8000000000 PRJNA704939
UWPOB_FLAV01	GCA_017302935.1	FLAV01	3.09	11,718	34	264	2	0	SAMN18059850 JAFB8000000000 PRJNA704939

(Continued on next page)

Volume 10 Issue 18 e00244-21

Isolate name	Assembly accession no.	Code	GTDB classification	Completeness (%)	Contamination (%)	Genome size (mbp)	No. of contigs	Avg contig size (bp)	% GC content	No. of rRNAs	No. of tRNAs	Relative abundance on date (mo/day/yr)					BioSample accession no.	GenBank accession no.	BioProject accession no.	
												7/16/2015	8/6/2015	7/24/2015	8/6/2015	Abundance				
JWPOB_FLAV02	GCA_017303755.1	FLAV02	d__Bacteri__Bacteroida; c__Bacteroidia__Flavobacteriales; f__Flavobacteriaceae; g__Flavobacterium__s__ d__Bacteri__Bacteroida; c__Bacteroidia__Flavobacteriales; f__Weeksiellaceae; g__Chryseobacterium__A__ s__Chryseobacterium__A__ d__Bacteri__Bacteroida; c__Ignavibacteri__Ignavibacteriales; f__Ignavibacteriaceae__Aug__UTCHB3__s__	100	0	4.07	15	271,271	34.7	1	46	1	0	0	0.44824	0.09632	0.2425	SAMN18059849	JAFLEB000000000	PRJNA704939
JWPOB_FLAV03	GCA_017303695.1	FLAV03	d__Bacteri__Bacteroida; c__Bacteroidia__Flavobacteriales; f__Weeksiellaceae; g__Chryseobacterium__A__ s__Chryseobacterium__A__ d__Bacteri__Bacteroida; c__Ignavibacteri__Ignavibacteriales; f__Ignavibacteriaceae__Aug__UTCHB3__s__	100	0	2.57	12	214,243	36.8	3	37	2	0	1	0.38721	0.07755	0.19068	SAMN18059851	JAFLEB000000000	PRJNA704939
JWPOB_IGNAV11	GCA_017302895.1	IGNAV11	d__Bacteri__Bacteroida; c__Bacteroidia__Flavobacteriales; f__Weeksiellaceae; g__Chryseobacterium__A__ s__Chryseobacterium__A__ d__Bacteri__Bacteroida; c__Ignavibacteri__Ignavibacteriales; f__Ignavibacteriaceae__Aug__UTCHB3__s__	100	0	3.98	6	663,902	42.2	3	44	1	1	1	2.23238	0.35053	0.08219	SAMN18059852	JAFLEB000000000	PRJNA704939
JWPOB_IGNAV12	GCA_017302885.1	IGNAV12	d__Bacteri__Bacteroida; c__Ignavibacteri__SJA-28f__g__s__	95.8	2.1	4.73	460	10,272	35.6	4	51	2	1	1	0.0987	0.13661	0.18405	SAMN18059853	JAFLEB000000000	PRJNA704939
JWPOB_IGNAV13	GCA_017303675.1	IGNAV13	d__Bacteri__Bacteroida; c__Ignavibacteri__SJA-28f__OLB5__g__OLB5__s__	93.1	0.6	3.65	3	1,217,215	37.8	3	43	1	1	1	13.6012	18.6822	18.7649	SAMN18059854	JAFLEB000000000	PRJNA704939
JWPOB_KAPA1	GCA_017302855.1	KAPA1	d__Bacteri__Bacteroida; c__Kapaobacteri__Kapaobacteriales; f__UBA961f__UBA2353__s__	93.5	0	3.6	309	11,645	38.6	0	34	0	0	0	0.0477	0.02835	0.0537	SAMN18059855	JAFLEB000000000	PRJNA704939
JWPOB_RHOD01	GCA_017303715.1	RHOD01	d__Bacteri__Bacteroida; c__Rhodothemia__Rhodothemiales; f__UBA2364g__UBA2364__s__	100	0	4.35	62	70,212	46.1	3	42	1	1	1	1.30416	0.50553	1.2655	SAMN18059856	JAFLEB000000000	PRJNA704939
JWPOB_DELTA1	GCA_017302875.1	DELTA1	d__Bacteri__Bacteroida; c__UBA2361f__UBA2361f__g__s__	95.7	3.8	3.86	296	13,034	52.6	3	45	1	1	1	0.16581	0.1972	0.0576	SAMN18059857	JAFLEB000000000	PRJNA704939
JWPOB_DELTA2	GCA_017302795.1	DELTA2	d__Bacteri__Bacteroida; c__Bacteri__Bacteroida; f__Bacteri__Bacteroida; g__Bacteri__Bacteroida; h__Bacteri__Bacteroida; i__Bacteri__Bacteroida; j__Bacteri__Bacteroida; k__Bacteri__Bacteroida; l__Bacteri__Bacteroida; m__Bacteri__Bacteroida; n__Bacteri__Bacteroida; o__Bacteri__Bacteroida; p__Bacteri__Bacteroida; q__Bacteri__Bacteroida; r__Bacteri__Bacteroida; s__	100	0	4.23	39	108,584	49.2	3	39	1	1	1	0.02329	0.29036	2.53023	SAMN18059858	JAFLEB000000000	PRJNA704939
JWPOB_DELTA3	GCA_017302835.1	DELTA3	d__Bacteri__Bacteroida; c__Bacteri__Bacteroida; f__Bacteri__Bacteroida; g__Bacteri__Bacteroida; h__Bacteri__Bacteroida; i__Bacteri__Bacteroida; j__Bacteri__Bacteroida; k__Bacteri__Bacteroida; l__Bacteri__Bacteroida; m__Bacteri__Bacteroida; n__Bacteri__Bacteroida; o__Bacteri__Bacteroida; p__Bacteri__Bacteroida; q__Bacteri__Bacteroida; r__Bacteri__Bacteroida; s__	99.1	0.9	3.63	35	103,678	37.3	3	32	1	1	1	0.73	0.4205	0.04244	SAMN18059859	JAFLEB000000000	PRJNA704939
JWPOB_DELTA4	GCA_017302715.1	DELTA4	d__Bacteri__Bacteroida; c__UBA2394g__UBA2394; f__UBA2394g__s__	100	0	2.39	21	113,698	41.4	4	41	1	1	2	0.09683	0.07426	0.01366	SAMN18059860	JAFLEB000000000	PRJNA704939
JWPOB_ANAER1	GCA_017303655.1	ANAER1	d__Bacteri__Bacteroida; c__Anaerolinea__Anaerolineales; f__envOPS12g__OLB14s__s__	83.1	6	4.61	172	26,782	49.6	1	40	1	0	0	0.78326	0.58803	0.33592	SAMN18059861	JAFLEB000000000	PRJNA704939
JWPOB_ANAER2	GCA_017303735.1	ANAER2	d__Bacteri__Bacteroida; c__Anaerolinea__Anaerolineales; f__envOPS12g__OLB14s__s__	87.5	6.4	4.8	421	11,388	50	0	47	0	0	0	0.46474	0.39027	0.53912	SAMN18059862	JAFLEB000000000	PRJNA704939
JWPOB_ANAER3	GCA_017303775.1	ANAER3	d__Bacteri__Bacteroida; c__Anaerolinea__SBR1031f__A4b; g__s__	85.2	0	4.47	499	8,952	56	2	29	2	0	0	0.22893	0.18326	0.09188	SAMN18059864	JAFLEB000000000	PRJNA704939
JWPOB_ANAER4	GCA_017303635.1	ANAER4	d__Bacteri__Bacteroida; c__Anaerolinea__SBR1031f__A4b; g__s__	99.5	0	6.8	39	174,361	56.3	3	48	1	1	1	0.50118	0.47335	0.26963	SAMN18059863	JAFLEB000000000	PRJNA704939
JWPOB_ANAER5	GCA_017304075.1	ANAER5	d__Bacteri__Bacteroida; c__Anaerolinea__SBR1031f__A4b; g__s__	91.7	5.5	7.11	706	10,069	60.1	2	45	2	0	0	0.16356	0.17579	0.09332	SAMN18059865	JAFLEB000000000	PRJNA704939
JWPOB_LEPT01	GCA_017302695.1	LEPT01	d__Bacteri__Cyanobacteria; c__Cyanobacteria__Leptolyngbyales; f__Leptolyngbyaceae; g__Leptolyngbya__Leptolyngbya	100	0	6.8	61	111,481	46.9	0	54	0	0	0	0.57222	0.82313	0.90187	SAMN18059866	JAFLEB000000000	PRJNA704939

mra.asm.org 3

Volume 10 Issue 18 e00244-21

Isolate name	Assembly accession no.	Code	GTDB classification	Completeness (%)	Contamination (%)	Genome size (mbp)	No. of contigs	Avg. contig size (bp)	% GC content	No. of rRNAs	No. of 5S rRNAs	No. of 16S rRNAs	Relative abundance on date (mo/day/yr)				Avg abundance	BioSample accession no.	GenBank accession no.	BioProject accession no.	
													7/16/2015	7/24/2015	8/6/2015	11/05/21					
JWPOB_OB51	GCA_017303615.1	OB51	d__Bacteri__Cyanobacteria; c__Vampirovibrionia; o__Obscuribacteriales; f__Obscuribacteraceae; g__Obscuribacter; d__Bacteri__Mycrococota; c__Polyangia__Polyangiales; f__Sindaracinaeae; s__	96.6	0.9	5.06	70	72,324	49	3	42	1	1	0.22519	0.60208	1.10521	0.64416	SAMN18059867	JAFLC0000000000	PRJNA704939	
JWPOB_DELTAS	GCA_017303575.1	DELTA5	d__Bacteri__Mycrococota; c__Polyangia__Polyangiales; f__Sindaracinaeae; s__	95.8	0	9.39	540	17,383	70.1	3	71	1	1	0.04541	0.10081	0.34331	0.16318	SAMN18059868	JAFLC0000000000	PRJNA704939	
JWPOB_PLANC1	GCA_017303565.1	PLANC1	d__Bacteri__Planctomycetota; c__Phycisphaerae__Phycisphaerales; f__SM1A02g__s__	88.3	6.3	2.93	530	5,521	68.6	1	37	0	1	0.02712	0.0326	0.0993	0.053	SAMN18059869	JAFLC0000000000	PRJNA704939	
JWPOB_PLANC2	GCA_017303555.1	PLANC2	d__Bacteri__Planctomycetota; c__Phycisphaerae__Phycisphaerales; f__SM1A02g__UBA2402s__	95.8	0	4.06	42	96,645	63.5	3	48	1	1	0.35589	0.44712	0.11479	0.30593	SAMN18059870	JAFLCN0000000000	PRJNA704939	
JWPOB_PLAC3	GCA_017303535.1	PLANC3	d__Bacteri__Planctomycetota; c__Planctomycetosa__Pirellulales; f__s__	93.5	1.3	6.61	641	10,313	61.8	3	71	1	1	0.02337	0.03888	0.08353	0.0486	SAMN18059871	JAFLC0000000000	PRJNA704939	
JWPOB_PLANC4	GCA_017303505.1	PLANC4	d__Bacteri__Planctomycetota; c__Planctomycetosa__Pirellulales; f__Pirellulaceae__Pirellula_Bs__	96.4	0.2	7.83	308	25,424	50.3	0	98	0	0	0.02843	0.10226	0.04436	0.05835	SAMN18059872	JAFLC0000000000	PRJNA704939	
JWPOB_PLANC5	GCA_017302815.1	PLANC5	d__Bacteri__Planctomycetota; c__UBA11346o__UBA11346i; f__UBA11346g__UBA11346s__	78	2	4.28	568	7,540	69.7	0	74	0	0	0.14524	0.27428	0.04468	0.15473	SAMN18059873	JAFLC0000000000	PRJNA704939	
JWPOB_ALPHA1	GCA_017302755.1	ALPHA1	d__Bacteri__Proteobacteria; c__Alphaproteobacteria__f__g__s__	94.8	0	2.12	275	7,722	59.4	2	35	1	0	1	0.00365	0.0307	0.09852	0.04429	SAMN18059874	JAFLC0000000000	PRJNA704939
JWPOB_CAULO1	GCA_017302335.1	CAULO1	d__Bacteri__Proteobacteria; c__Alphaproteobacteria; o__Caulobacteriales; f__Caulobacteraceae; g__Brevundimonas; s__	96.6	0.7	2.6	168	15,478	67.9	1	44	0	1	0	0.29751	0.15141	0.02773	0.15888	SAMN18059876	JAFLC0000000000	PRJNA704939
JWPOB_CAULO2	GCA_017302105.1	CAULO2	d__Bacteri__Proteobacteria; c__Alphaproteobacteria; o__Caulobacteriales; f__Caulobacteraceae; g__Brevundimonas; s__	100	0	2.57	12	214,253	65.7	3	38	1	1	1.58091	0.90968	0.11758	0.86939	SAMN18059875	JAFLC0000000000	PRJNA704939	
JWPOB_CAULO3	GCA_017303355.1	CAULO3	d__Bacteri__Proteobacteria; c__Alphaproteobacteria; o__Caulobacteriales; f__Caulobacteraceae; g__Brevundimonas; s__	98.4	0.7	3.67	34	107,997	63.9	3	47	1	1	0.9416	0.71847	0.41979	0.69328	SAMN18059877	JAFLC0000000000	PRJNA704939	
JWPOB_CAULO4	GCA_017303435.1	CAULO4	d__Bacteri__Proteobacteria; c__Alphaproteobacteria; o__Caulobacteriales; f__Hyphomonadaceae__SWB02s__	75	6.3	3.49	362	9,650	39.2	0	35	0	0	0.09806	0.03402	0.00279	0.04496	SAMN18059878	JAFLCV0000000000	PRJNA704939	
JWPOB_MICAV1	GCA_017302165.1	MICAV1	d__Bacteri__Proteobacteria; c__Alphaproteobacteria; o__Micavibrionales; f__Micavibrionaceae; s__	96	0.2	2	6	333,677	49.3	2	40	0	1	1	0.02039	0.01146	0.35063	0.1275	SAMN18059879	JAFLCV0000000000	PRJNA704939
JWPOB_RHZO1	GCA_017302635.1	RHZO1	d__Bacteri__Proteobacteria; c__Alphaproteobacteria; o__Rhizobiales__Beijerinckaceae_A; g__PAR1s__	97.3	0.6	3.1	6	516,587	62.1	3	47	1	1	0.21701	0.24399	0.35215	0.27105	SAMN18059880	JAFLC0000000000	PRJNA704939	
JWPOB_RHODOB1	GCA_017303485.1	RHODOB1	d__Bacteri__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae; g__Tabrizicola; s__	96.6	8.6	4.39	272	16,121	65.3	0	53	0	0	0.06267	0.10359	0.03424	0.06683	SAMN18059881	JAFLCV0000000000	PRJNA704939	
JWPOB_RICK1	GCA_017302665.1	RICK1	d__Bacteri__Proteobacteria; c__Alphaproteobacteria; o__Rickettsiales__Rickettsiaceae; s__	98.6	7.1	1.65	29	56,773	33.3	4	34	1	1	2	0.00855	0.16939	0.03581	0.07125	SAMN18059882	JAFLC2000000000	PRJNA704939

mra.asm.org 4

TABLE 1 (Continued)

Isolate name	Assembly accession no.	Code	GTDB classification	Complete-ness (%)	Contami-nation (%)	Genome size (mbp)	No. of contigs	Avg. contig size (bp)	% GC content	No. of rRNAs	No. of tRNAs	No. of 5S rRNAs	No. of 16S rRNAs	No. of 23S rRNAs	Relative abundance on date (mol/day/yr)			Avg abundance	BioSample accession no.	GenBank accession no.	BioProject accession no.
															7/16/2015	7/24/2015	8/6/2015				
UWPOB_RICK2	GCA_017302595.1	RICK2	d__Bacteriaph__Proteobacteria; c__Alphaproteobacteria; o__Rickettsiales; f__Rickettsiaceae; g__GCA_24021955	99.5	1	1.33	5	264,993	34.7	0	34	0	0	0	0.03107	0.20695	0.08658	0.1082	SAMN18059883	JAFLEDA0000000000	PRJNA704939
UWPOB_SPHING1	GCA_017302775.1	SPHING1	d__Bacteriaph__Proteobacteria; g__GCA_24021955	78.5	5.1	3.16	556	5,684	61.6	0	48	0	0	0	0.06166	0.0599	0.01391	0.04516	SAMN18059885	JAFLEDC0000000000	PRJNA704939
UWPOB_SPHING2	GCA_017302615.1	SPHING2	d__Bacteriaph__Proteobacteria; o__Sphingomonadales; f__Sphingomonadaceae; g__Novosphingobium	99.2	0.1	3.02	33	91,486	65.7	3	46	1	1	1	0.6247	0.62665	0.23261	0.49465	SAMN18059884	JAFLEDB0000000000	PRJNA704939
UWPOB_ALPHA2	GCA_017302575.1	ALPHA2	d__Bacteriaph__Proteobacteria; c__Alphaproteobacteria; o__Sphingomonadales; f__Sphingomonadaceae; g__Novosphingobium	98.9	0	1.86	1	1,861,815	50.8	3	39	1	1	1	0.47594	0.19689	0.04249	0.23844	SAMN18059886	JAFLEDD0000000000	PRJNA704939
UWPOB_BURK1	GCA_017302655.1	BURK1	d__Bacteriaph__Proteobacteria; f__UBA3002g_s	100	0	4.44	203	21,857	67.3	2	48	1	1	0	0.37852	0.42244	0.55223	0.45106	SAMN18059887	JAFLEDE0000000000	PRJNA704939
UWPOB_OTTO1	GCA_017302725.1	OTTO1	d__Bacteriaph__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__s	95.8	0	2.87	325	8,835	68.6	3	52	1	1	1	0.76562	0.83948	0.30961	0.63823	SAMN18059888	JAFLEDF0000000000	PRJNA704939
UWPOB_RUBR1	GCA_017302495.1	RUBR1	d__Bacteriaph__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__Rubrivivax	81	0.8	4.51	553	8,152	71.1	3	56	1	1	1	1.62184	1.49067	0.45613	1.18955	SAMN18059889	JAFLEDG0000000000	PRJNA704939
UWPOB_RUBR2	GCA_017302505.1	RUBR2	d__Bacteriaph__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__Rubrivivax	100	0	5.5	555	9,913	68.4	4	52	2	1	1	0.21114	0.24105	0.10683	0.18634	SAMN18059890	JAFLEDH0000000000	PRJNA704939
UWPOB_VITRI1	GCA_017302485.1	VITRI1	d__Bacteriaph__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__Vitreoscilla	95.8	0	3.68	513	7,177	68.4	1	42	1	0	0	0.07383	0.05699	0.17722	0.10268	SAMN18059891	JAFLEDI0000000000	PRJNA704939
UWPOB_BURK2	GCA_017302475.1	BURK2	d__Bacteriaph__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Palka-1005g_s	79.2	0	2.71	498	5,448	69.6	0	34	0	0	0	0.10158	0.12158	0.11145	0.11154	SAMN18059892	JAFLEDJ0000000000	PRJNA704939
UW10-POB	GCA_017302555.1	CAPIB	d__Bacteriaph__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Accumulibacter	100	0	4.4	88	49,949	62.5	0	41	0	0	0	0.52449	0.06097	0.65232	0.59459	SAMN18059893	JAFLEDK0000000000	PRJNA704939
UW12-POB	GCA_017302435.1	CAPID	d__Bacteriaph__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Accumulibacter	98.1	0.7	4.56	91	50,155	62.7	0	46	0	0	0	0.35414	0.32156	0.05167	0.24246	SAMN18059894	JAFLEDL0000000000	PRJNA704939
UW13-POB	GCA_017302415.1	CAPIF	d__Bacteriaph__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Accumulibacter	97.1	4.4	5.45	245	22,232	65.9	1	48	0	1	0	1.65004	0.70231	0.09466	0.81567	SAMN18059895	JAFLEDM0000000000	PRJNA704939

(Continued on next page)

Volume 10 Issue 18 e00244-21

Isolate name	Assembly accession no.	Code	GTDB classification	Completeness (%)	Contamination (%)	Genome size (mbp)	No. of contigs	Avg. contig size (bp)	% GC content	No. of rRNAs	No. of tRNAs	No. of 5S rRNAs	No. of 16S rRNAs	Relative abundance on date (mo/day/yr)					Avg. abundance	BioSample accession no.	GenBank accession no.	BioProject accession no.	
														7/16/2015	7/24/2015	8/6/2015	10/23/2015	11/4/2015					
JW11-POB	GCA_017302385.1	CAPIC	d__Bacteriopl__Proteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Accumulibacter__Accumulibacter	100	0	4.63	144	32,127	61.2	0	47	0	0	12.2117	11.8669	10.2307	11.4364	SAMN18059898	JAFLDP0000000000	PRJNA704939			
JW8-POB	GCA_017302345.1	CAPIA	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Accumulibacter__Accumulibacter	100	0	4.26	49	86,993	64.2	0	51	0	0	6.34328	6.19256	0.4336	4.32315	SAMN18059896	JAFLDN0000000000	PRJNA704939			
JW9-POB	GCA_017302455.1	CAPIA	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Accumulibacter__Accumulibacter	100	0	4.93	126	39,138	64.1	1	46	1	0	0.55782	0.54654	0.15933	0.42123	SAMN18059897	JAFLDQ0000000000	PRJNA704939			
JWPOB_DECH	GCA_017302355.1	DECH1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Accumulibacter__Accumulibacter	91.5	8.1	3.69	279	13,228	65	2	40	0	1	2.09698	3.33708	3.17482	2.86963	SAMN18059899	JAFLDQ0000000000	PRJNA704939			
JWPOB_SULF1	GCA_017302275.1	SULF1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Dechloromonas	98.3	1.1	4.41	88	50,106	64	3	47	2	1	0.76164	0.92706	2.2433	1.31067	SAMN18059900	JAFLDR0000000000	PRJNA704939			
JWPOB_THAU1	GCA_017302295.1	THAU1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Sulfuritalea	98.7	1.4	4.11	56	73,417	66.1	3	52	3	0	0.5635	0.57917	0.40828	0.51699	SAMN18059901	JAFLD5000000000	PRJNA704939			
JWPOB_ZOOG1	GCA_017309145.1	ZOOG1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Thauera	83.3	4.2	4.91	428	11,461	65.7	0	58	0	0	0.06958	0.06569	0.07688	0.07071	SAMN18059903	JAFLDU0000000000	PRJNA704939			
JWPOB_ZOOG2	GCA_017302315.1	ZOOG2	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Burkholderiales; f__Rhodocyclaceae; g__Zoogloea	95.2	1.8	4.52	165	27,407	64.5	0	55	0	0	0.12988	0.14704	0.21305	0.16316	SAMN18059902	JAFLDT0000000000	PRJNA704939			
JWPOB_PARA1	GCA_017302255.1	PARA1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Alteromonadaceae; g__Parahelminthina	100	4.2	4.03	77	52,358	46.2	0	50	0	0	0.20334	0.02657	0.00492	0.07828	SAMN18059904	JAFLDV0000000000	PRJNA704939			
JWPOB_GAMMA1	GCA_017302205.1	GAMMA1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Xanthomonadales; f__Rhodocyclaceae; g__Dokdonella	79.2	1	2.51	456	5,508	69.2	1	34	1	0	0.05103	0.07089	0.10094	0.07429	SAMN18059905	JAFLDW0000000000	PRJNA704939			
JWPOB_RHODAN1	GCA_017302215.1	RHODAN1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Xanthomonadales; f__Rhodanobacteraceae; g__Xanthomonas	91.7	0	3.5	323	10,843	68.8	0	49	0	0	0.144	0.20595	0.16775	0.17257	SAMN18059906	JAFLDX0000000000	PRJNA704939			
JWPOB_AQU1	GCA_017302155.1	AQU1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Xanthomonadales; f__Xanthomonadaceae; g__Aquimonas	98.3	1.6	3.98	113	35,212	61.5	0	46	0	0	0.28393	0.24393	0.06593	0.19793	SAMN18059907	JAFLDY0000000000	PRJNA704939			
JWPOB_LYSO1	GCA_017302135.1	LYSO1	d__Bacteriopl__Proteobacteria; c__Gammaproteobacteria; o__Xanthomonadales; f__Xanthomonadaceae; g__Lysobacter	99.6	1	4.34	35	123,959	64.1	2	48	1	1	0.121091	1.20058	0.96205	1.12451	SAMN18059908	JAFLDZ0000000000	PRJNA704939			

mra.asm.org 6

TABLE 1 (Continued)

Isolate name	Assembly accession no.	Code	GTDB classification	Completeness (%)	Contamination (%)	Genome size (mbp)	No. of contigs	Avg contig size (bp)	% GC content	No. of rRNAs	No. of tRNAs	No. of 5S rRNAs	No. of 16S rRNAs	No. of 23S rRNAs	Relative abundance on date (mo/day/yr)			Avg abundance	BioSample accession no.	GenBank accession no.	BioProject accession no.
															7/16/2015	7/24/2015	8/6/2015				
UWPOB_PSEUDO1	GCA_017302035.1	PSEUDO1	d_Bacteriap__Proteobacteria; c_Gammaproteobacteria; o_Xanthomonadales; f_Xanthomonadaceae; g_Pseudoxanthomonas; d_Bacteriap__Proteobacteria; c_Gammaproteobacteria; o_Xanthomonadales; f_Xanthomonadaceae; g_Thermomonas	95.8	0	2.91	125	23,300	69.3	1	48	1	0	0	1.50336	2.17593	0.47299	1.38409	SAMN18059909	JAFLEA0000000000	PRJNA704939
UWPOB_THERMO1	GCA_017302035.1	THERMO1	d_Bacteriap__Proteobacteria; c_Gammaproteobacteria; o_Xanthomonadales; f_Xanthomonadaceae; g_Thermomonas	76.4	2.1	2.31	299	7,734	69.7	0	39	0	0	0	0.21965	0.25626	0.05906	0.17832	SAMN18059911	JAFLEA0000000000	PRJNA704939
UWPOB_THERMO2	GCA_017302035.1	THERMO2	d_Bacteriap__Proteobacteria; c_Gammaproteobacteria; o_Xanthomonadales; f_Xanthomonadaceae; g_Thermomonas	87.5	0.5	2.77	245	11,303	69.9	0	41	0	0	0	0.79368	1.14049	0.4586	0.79759	SAMN18059910	JAFLEA0000000000	PRJNA704939
UWPOB_SPIRO1	GCA_017302045.1	SPIRO1	d_Bacteriap__Spirochaetota; c_Leptospirae; f_Turnelliales; g_Turnelliales; d_Bacteriap__Spirochaetota; c_UBA121350__f_g_s	86.5	0	3.78	412	9,184	53.8	2	31	2	0	0	0.03758	0.03209	0.0621	0.04392	SAMN18059912	JAFLEA0000000000	PRJNA704939
UWPOB_SPIRO2	GCA_017302015.1	SPIRO2	d_Bacteriap__Spirochaetota; c_UBA121350__f_g_s	87.5	0	5.43	173	31,368	65.2	2	41	0	1	1	0.09083	0.09013	0.07872	0.08656	SAMN18059913	JAFLEA0000000000	PRJNA704939
UWPOB_VERRUCO1	GCA_017302195.1	VERRUCO1	d_Bacteriap__Verrucomicrobia; c_Verrucomicrobiae; o_Pedospherales; f_UBA9464; g_UBA9464s	81	8.3	4.86	848	5,726	67.1	0	35	0	0	0	0.04712	0.07195	0.01614	0.04507	SAMN18059915	JAFLEA0000000000	PRJNA704939
UWPOB_VERRUCO2	GCA_017309115.1	VERRUCO2	d_Bacteriap__Verrucomicrobia; c_Verrucomicrobiae; o_Pedospherales; f_UBA9464; g_UBA9464s	98.6	3.5	7.53	112	67,228	61.5	3	50	1	1	1	0.14886	0.18964	0.1474	0.16196	SAMN18059914	JAFLEA0000000000	PRJNA704939
UWPOB_VERRUCO3	GCA_017301995.1	VERRUCO3	d_Bacteriap__Verrucomicrobia; c_Verrucomicrobiae; o_Verrucomicrobiales; f_Akkermansia; g_UBA1315s	100	0	7.04	68	103,585	62.3	2	60	2	0	0	0.30515	0.26629	0.07792	0.21646	SAMN18059916	JAFLEA0000000000	PRJNA704939
UWPOB_VERRUCO4	GCA_017301975.1	VERRUCO4	d_Bacteriap__Verrucomicrobia; c_Verrucomicrobiae; o_Verrucomicrobiales; f_Verrucomicrobiae; g_Verrucomicrobiae	97	0	6.9	97	71,172	60.3	0	54	0	0	0	0.17845	0.15859	0.07204	0.13636	SAMN18059917	JAFLEA0000000000	PRJNA704939

^a Genome information for 86 metagenome-assembled genomes is shown. The column labeled "Isolate name" lists the name submitted to NCBI, and that labeled "Code" lists the codes used in Fig. 1. Classifications were assigned with GTDB-tk and "Candidatus Accumulibacter phosphatis" clade designations based on comparisons to *ppk1* clones and publicly available reference genome sequences. "Candidatus Accumulibacter phosphatis" genomes are named alphanumerically according to previous nomenclature; all other genomes are named UWPOB_CODE. The genome statistics for completeness, contamination, GC content, genome size, and number of contigs were performed with CheckM v1.1.2. The rRNA and tRNA genes were predicted with Barrnap v0.9, as part of the Prokka v1.13.7 package. The relative abundance was calculated using coverM v0.4.0 using the relative abundance method for the three metagenomic samples, taken on 16 July 2015, 24 July 2015, and 6 August 2015 toward the end of the enrichment period, and averaged together across all three samples for an average relative abundance calculation.

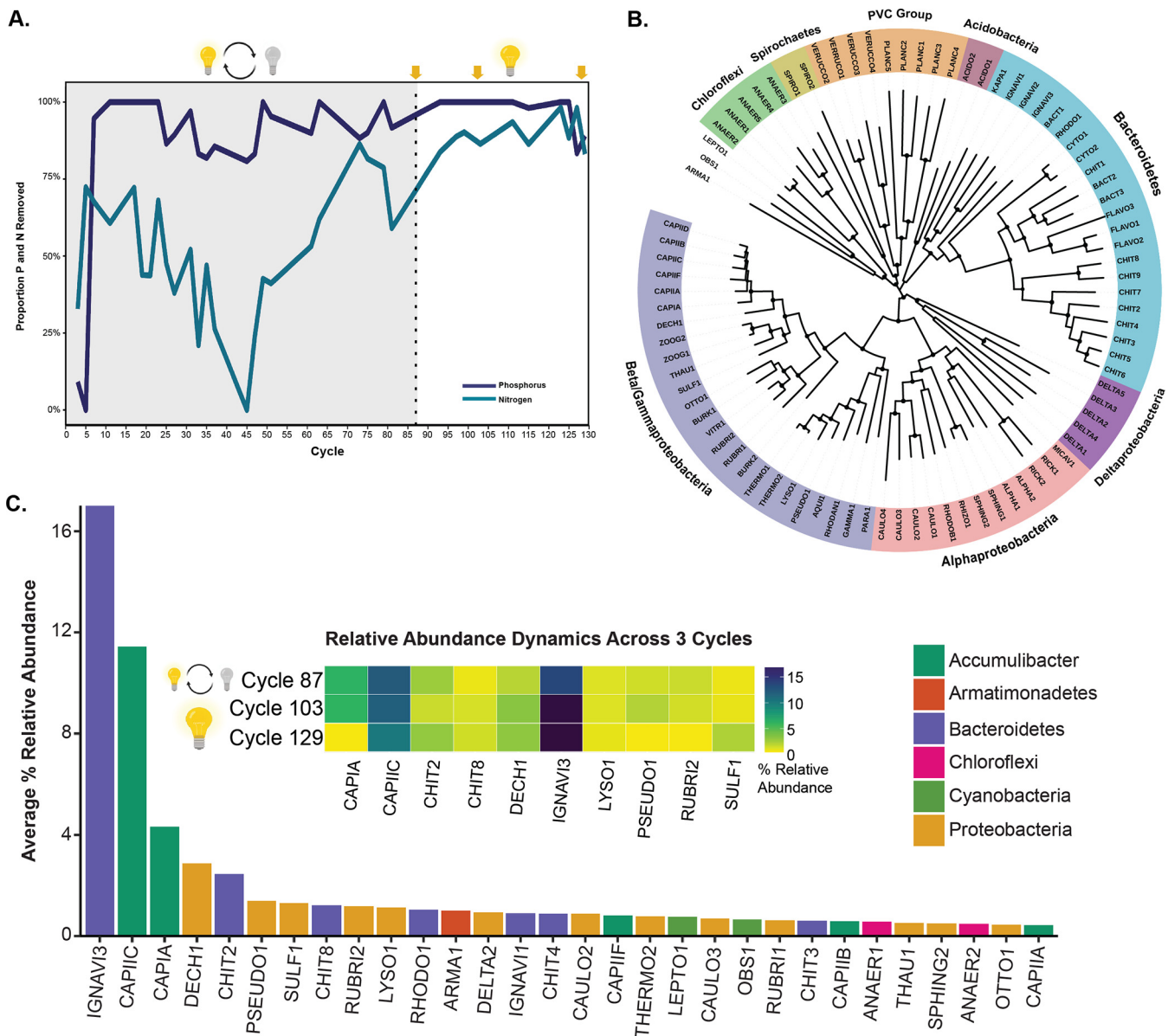


FIG 1 Overview of reactor performance and assembled MAGs. (A) Proportions of nitrogen (blue) and phosphorus (purple) removed during reactor operation. Cycle length refers to the hydraulic residence time of 0.625 days (2). The reactor was operated in two phases—in the first phase (gray), light was cycled on and off, and in the second phase (white), light was provided continuously. The arrows denote the time points at which metagenomic sequencing was performed. (B) Phylogenetic tree of all 86 assembled MAGs constructed from a concatenated alignment of 120 single-copy marker genes from GTDB-tk v0.3.2 (9) that were identified and aligned using HMMER v3.2.1 (11). The tree was constructed using RAXML v8.2.12 with 100 rapid bootstraps (12) and viewed using iTOL v5.7 (13). (C) Relative abundance and population dynamics of abundant species represented by assembled MAGs. The bar graph represents the average relative abundance of the top 30 most abundant species across all three samples. Relative abundance calculations were made using coverM v0.4.0 with the relative_abundance method and averaged across all three samples. The inset heat map shows the population dynamics of the top 10 most abundant species, representing the relative abundance of that MAG in that sample.

sequences were checked for uniform differential coverage using uBin v0.9.14 (7). All genome statistics were calculated with CheckM v1.1.2 (8). Taxonomical classifications were made based on the Genome Taxonomy Database (GTDB) using GTDB-Tk v0.3.2 (9, 10). Relative abundance calculations were performed with coverM v0.4.0 using the relative_abundance calculation method. The methods and phylogenetic tree for assigning clades of six “*Ca. Accumulibacter phosphatis*” draft genome sequences based on the *ppk1* locus are available at <https://doi.org/10.6084/m9.figshare.14164478>.

Data availability. The raw metagenomes for the 3 samples and genome assemblies for all 86 metagenome-assembled genomes (MAGs) are available at NCBI GenBank

under BioProject accession number [PRJNA704939](#). The metagenomes are available under the SRA accession numbers [SRR13786854](#), [SRR13786855](#), and [SRR13786856](#).

ACKNOWLEDGMENTS

This work was supported by funding from the National Science Foundation (MCB-1518130) to K.D.M. and D.R.N.

We thank the University of Wisconsin—Madison Biotechnology Center DNA Sequencing Facility for Illumina sequencing services. This research was performed in part using the Wisconsin Energy Institute computing cluster, which is supported by the Great Lakes Bioenergy Research Center as part of the U.S. Department of Energy Office of Science (DE-SC0018409).

REFERENCES

1. Martin HG, Ivanova N, Kunin V, Warnecke F, Barry KW, McHardy AC, Yeates C, He S, Salamov AA, Szeto E, Dalin E, Putnam NH, Shapiro HJ, Pangilinan JL, Rigoutsos I, Kyripides NC, Blackall LL, McMahon KD, Hugenholtz P. 2006. Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. *Nat Biotechnol* 24:1263–1269. <https://doi.org/10.1038/nbt1247>.
2. Oyserman BO, Martirano JM, Wipperfurth S, Owen BR, Noguera DR, McMahon KD. 2017. Community assembly and ecology of activated sludge under photo-synthetic feast-famine conditions. *Environ Sci Technol* 51:3165–3175. <https://doi.org/10.1021/acs.est.6b03976>.
3. Bushnell B, Rood J, Singer E. 2017. BBMerge—accurate paired shotgun read merging via overlap. *PLoS One* 12:e0185056. <https://doi.org/10.1371/journal.pone.0185056>.
4. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
5. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3:e1165. <https://doi.org/10.7717/peerj.1165>.
6. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 11:2864–2868. <https://doi.org/10.1038/ismej.2017.126>.
7. Bornemann TLV, Esser SP, Stach TL, Burg T, Probst AJ. 2020. uBin—a manual refining tool for metagenomic bins designed for educational purposes. *bioRxiv*. <https://doi.org/10.1101/2020.07.15.204776>.
8. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
9. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a tool-kit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>.
10. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 36:996–1004. <https://doi.org/10.1038/nbt.4229>.
11. Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol* 7:e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
12. Stamatakis A. 2014. The RAxML v8.0.X manual. <https://cme.h-its.org/exelixis/resource/download/NewManual.pdf>.
13. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:W242–W245. <https://doi.org/10.1093/nar/gkw290>.