ACS ES&T Engineering

Open Access

pubs.acs.org/estengg

Article

Reverse Osmosis in an Advanced Water Treatment Train Produces a Simple, Consistent Microbial Community

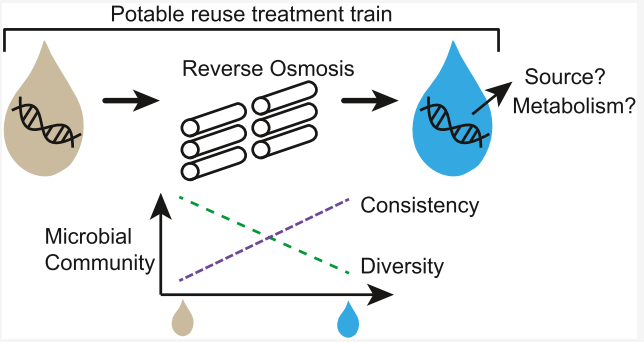
Rose S. Kantor,* Lauren C. Kennedy, Scott E. Miller, Jorien Favere, and Kara L. Nelson*

Cite This: ACS EST Engg. 2025, 5, 772−781

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Potable water reuse has become a key component of water sustainability planning in arid regions. Many advanced water purification facilities use reverse osmosis (RO) as part of treatment, including as a barrier for microorganisms; however, regrowth after RO treatment has been observed. Questions remain about the identity, source, and survival mechanisms of microorganisms in RO permeate, but the extremely low biomass of this water is a limitation for common microbiological methods. Here, we performed high-throughput sequencing on samples collected throughout a potable reuse train, including samples collected by filtering large volumes of RO permeate and biomass collected from RO membranes during an autopsy. We observed a stable, consistent microbial community across three months and in two parallel RO trains. RO permeate samples contained Burkholderiaceae at high relative abundance, including one *Aquabacterium* sp. that accounted for 29% of the community, on average. Like most other RO permeate microorganisms, this sequence was not seen in upstream samples and we suggest that biofilm growing on unit process infrastructure, rather than active treatment breakthrough, was the primary source. A metagenome-assembled genome corresponding to *Aquabacterium* sp. from RO permeate was found to lack most sugar-utilization pathways and to be able to consume low molecular weight organic molecules, potentially those that pass through RO.



KEYWORDS: *potable reuse, microbiology, reverse osmosis, metagenomics, 16S*

1. INTRODUCTION

Exacerbated by climate change, water scarcity is affecting more people worldwide.[1] In California, water agencies are seeking diversification of the water portfolio to meet supply needs. This water portfolio includes both potable reuse and desalination,[2] both of which rely on membrane-based processes for treatment. Potable reuse of wastewater via advanced treatment makes use of an existing, reliable, and often untapped supply of water and has been applied globally.[3−5] After advanced treatment, water for potable reuse can be sent to an aquifer or reservoir or directly to a drinking water treatment plant or distribution system.

Membrane processes like microfiltration (MF) and reverse osmosis (RO) are integral components of desalination and full advanced treatment of wastewater in California. Upstream steps sometimes included in potable reuse treatment trains include ozonation followed by biological activated carbon (BAC) filtration, and downstream of RO, permeate is further treated with advanced oxidation processes (AOP). Studies on the efficacy of advanced treatment trains primarily focus on robust removal of indicator viruses, bacteria, protozoan cysts, and chemical constituents of concern.[6−8] Regulations are developed to address these specific contaminants known to present human health risks, but bacterial cells are present following RO,[9] and few studies have investigated the origins of these microorganisms.

While high levels of removal and inactivation of pathogenic and indicator bacteria occur in advanced treatment trains, we have previously shown that bacteria can grow after these processes, with the extent of regrowth depending on local conditions.[9−12] Even if bacteria are removed and inactivated by RO and AOP, respectively, nutrients and carbon may remain, contributing to bacterial growth during water distribution and storage.[11] Of biological interest, RO permeate represents an extreme environment due to the low ionic strength and very low nutrient content.[13] Furthermore, RO membranes are expensive to operate, and fouling increases energy requirements and capital costs due to cleaning and replacement.[14]

Received: October 3, 2024
Revised: January 23, 2025
Accepted: January 24, 2025
Published: February 4, 2025

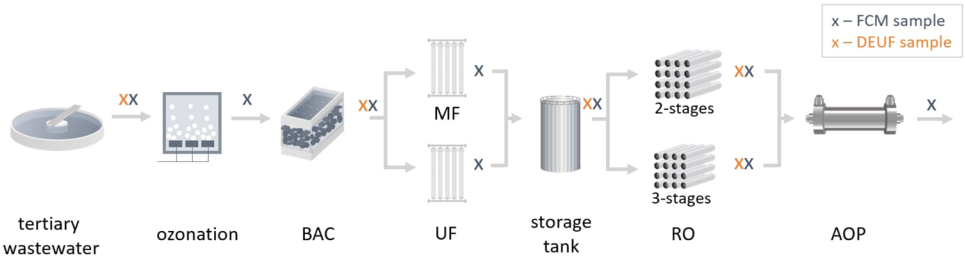


Figure 1. Overview of the treatment train. Tertiary treated wastewater was treated sequentially by ozonation, BAC filtration, parallel MF, and ultrafiltration (UF), parallel reverse osmosis units with two or three stages in series (RO), and a UV/free chlorine AOP. Sample points for flow cytometry (FCM) and dead-end ultrafiltration (DEUF) are indicated by gray and orange *x*'s, respectively.

This motivates investigation of the microorganisms found on the feed side of RO membranes.

Specifically, it could be beneficial to better understand the microbial biomass surrounding RO in the context of a potable reuse treatment train. Relatively few membrane-based treatment facilities have been studied from a microbiological perspective, and each has a unique treatment train and source water.[10,15−19] Studies on advanced treatment facility microbiomes[10,18] have been limited in their ability to obtain sequence data for bacteria in reverse osmosis permeate because the cell concentrations were very low. We previously investigated the microbial water quality within a demonstration-scale advanced treatment facility[9] during a period when chloramines were not added to the feed of the MF and RO,[20] potentially causing increased cell counts in the permeate and, thus, improved collection of biomass. This previous work reported on the log-removals of cells and bacterial genes across the treatment train and showed that, despite the lack of chloramine, very low biomass remained after RO.

Here, we applied two high-throughput sequencing techniques to these previously collected samples to address the following questions: (1) How does the diversity and consistency of the RO permeate microbial community compare to other treatment steps in the advanced treatment train? (2) What is the identity and potential source of microorganisms in RO permeate? (3) Is cell growth different after RO, and what metabolisms might contribute to growth? We utilized 16S rRNA gene amplicon sequencing for its sensitivity, and genome-resolved metagenomics to gain insight into metabolic potential. To our knowledge, this is the first study to perform sequencing on large sample volumes of RO permeate and biomass recovered from RO membranes in an advanced wastewater treatment facility.

2. METHODS

2.1. Treatment Train and Operation. The advanced wastewater treatment demonstration facility had a treatment capacity of approximately 3.8×10^6 L/day and was designed to test the performance of several parallel treatment processes (Figure 1).

The complete treatment train has been described by ref 9. Briefly, the influent was tertiary wastewater (treated by sedimentation, activated sludge, and granular media filtration with anthracite coal). In the advanced treatment demonstration facility, treatment included (i) ozonation (Wedeco/Xylem, DC, USA), with an average target applied ozone concentration of 8.17 mg/L, CT range of 1.84−4.00 (average = 2.68) mg-min/L, and hydraulic retention time of ∼7.5 min to target 1-$\log_{10}$ reduction of *Cryptosporidium*; (ii) two parallel, identical biologically active granular activated carbon filters ("BAC", Leopold/Xylem, DC, USA), operated with empty bed contact times of 15 min and average backwash intervals of approximately two to 4 days; (iii) parallel microfiltration ("MF", Pall Corporation, New York) and ultrafiltration ("UF", modules by Toray, Japan, and overall design by H_2O Innovation, Canada) membranes, with nominal pore sizes of 0.1 and 0.015 μm, respectively, and average water recoveries of 96% each; (iv) two parallel reverse osmosis ("RO", EnAqua, California) units, one with two stages in series and the other with three stages in series, each operated at an average water recovery of 75−80%; and (v) a UV and free chlorine advanced oxidation process ("AOP" system, TrojanUV, Canada), with free chlorine concentrations in the feed and effluent of approximately 2 and 1−1.5 mg/L as Cl_2, respectively, hydraulic retention time of approximately 15 s, and minimum UV dose of 850 mJ/cm^2.

2.2. Sampling. Sampling for flow cytometric cell counts and nucleic acids analysis was conducted 2−4 weeks apart between September and December 2017[9] (Table S1). Grab samples were collected for flow cytometry from all locations (Figure S1), with excess sodium thiosulfate added to quench free chlorine in AOP effluent. Flow cytometry methods and data are fully described.[9] Microbial biomass was concentrated by dead-end ultrafiltration (DEUF) from tertiary wastewater, BAC filtrate, MF/UF combined permeate storage tank (referred to as "MF" throughout), and RO permeates (Figure 1), following an adaptation of the method described previously.[21] In brief, a range of water volumes, depending on the sampling location (30 L for tertiary wastewater up to 4000 L for RO permeate, which has low biomass and requires collection of large volumes) was collected on ultrafiltration membranes (REXEED 25S, Henry Schein, New York) (Table S1). The filters were backflushed with 500 mL of a backflush solution (0.5% w/v Tween 80, 0.01% w/v sodium polyphosphate, and 0.001% w/v Y-30 antifoam emulsion) through the ultrafilter and into a sterile container in the opposite direction from sample filtering. The backflush underwent a second concentration step via poly(ethylene glycol) (PEG) precipitation. PEG pellets containing microbial biomass were then resuspended in ∼3 mL of supernatant or TE buffer and stored as two aliquots at −80 °C until DNA extraction. To capture potential contaminants from the filters and solutions used during sample processing, 1 L of autoclaved deionized (DI) water was filtered through clean ultrafiltration membranes and processed alongside samples as "Field Blanks".[9]

An RO membrane autopsy was conducted on January 10, 2018 (day 118) by the facility to observe membrane fouling. Our research team was invited to collect biofilm samples during this process. Once the membrane was opened (Figure

https://doi.org/10.1021/acsestengg.4c00665
ACS EST Engg. 2025, 5, 772−781

S2), biofilm from the feed side of the membrane was collected directly into a sterile tube using a sterile cell scraper. The mesh separator on the permeate side was cut with sterile scissors and placed into autoclaved DI water. A separate bottle of autoclaved DI water was placed on the sampling surface and opened during sample collection as a sampling negative control. Samples and controls were transported to the lab on ice and processed immediately. Retentate biofilm was stored at −80 C. The mesh separator in water and the negative control water were sonicated to wash the mesh and dislocate biofilm. The mesh was removed, and the wash water and negative control underwent PEG precipitation following the same protocol used for all ultrafiltration backflush samples. The facility conducted a second autopsy on August 8, 2018 and the feed and permeate biofilm samples were included in further processing for DNA extraction and sequencing and are shown in the SI for comparison. These samples are referred to as "ROBF" feed and permeate throughout.

2.3. DNA Extraction. DNA was extracted from sample pellets using a PowerSoil Pro extraction kit (Qiagen, Maryland) according to the manufacturer's protocol, with slight modifications. Briefly, PEG pellets were thawed and homogenized by vigorously vortexing for 10 s. For tertiary influent, BAC, MF/UF storage tank, and 4 RO permeate samples, 200 μL of homogenized sample was added directly to the PowerSoil Pro Powerbead tube. For the remaining RO samples, homogenized pellets from identical aliquots were centrifuged (15 000g for 15 min). The supernatant was combined onto a centrifugal filtration unit (Amicon ultra-15 centrifugal filter unit, 100 kDa; Millipore, Ireland) and centrifuged (7500g for 30 min). The concentrated supernatant was used to resuspend the centrifuged pellets, and 200 μL of the resulting sample was added to the Powerbead Tube. All samples were incubated at 37 C for 30 min in an enzymatic digestion solution: 50 μL of 0.001% lysozyme (Sigma-Aldrich, Germany), 50 μL of 0.00001% achromopeptidase (Sigma-Aldrich, Germany), and 8 μL of 0.01% carrier RNA in buffer AVL (Qiagen, Maryland). Finally, 500 μL of solution CD1 (PowerSoil Pro) was added, and extraction followed the PowerSoil Pro manufacturer protocol. The elution buffer was incubated on the column for 5 min at room temperature prior to the final elution step. Eluates were stored at −80 C. The effective volume extracted for each sample is reported in Table S1. Due to limited sample volumes, three samples that were part of initial methods testing experiments were included in the final set for sequencing. Of these, two were RO permeate samples extracted with the Qiagen DNeasy Blood and Tissue kit, while the other was a BAC filtrate sample extracted with the Qiagen DNeasy PowerMax soil kit followed by ethanol precipitation (Table S1). The Zymobiomics mock microbial community standard was used for seven positive extraction controls spanning 6 orders of magnitude in estimated cell inputs; three of these were sequenced by metagenomics (below) and while all were included in amplicon sequencing. Nuclease-free water was used as a negative control in each DNA extraction batch (n = 6).

2.4. Library Preparation and Sequencing. Amplicon library preparation for the V4 region of the 16S rRNA gene was performed as part of a larger project previously described.[11] Briefly, library preparation followed the Schloss Lab MiSeq wet-lab protocol,[22] using dual barcoded primers (515F and 806R) in a one-step PCR amplification reaction for each sample. Triplicate 25 μL PCR reactions were combined prior to cleanup and barcoded amplicons were pooled based on Qubit values prior to sequencing. Controls included no-template negative PCR controls on each plate (n = 2), and the Zymobiomics mock microbial community DNA standard (positive PCR control). Sequencing was performed with Illumina MiSeq (v3), yielding 300 bp paired-end reads.

Library preparation for metagenomic sequencing was performed at the Functional Genomics Laboratory (UC Berkeley). Briefly, 500 bp insert libraries were constructed using the KAPA HyperPrep kit with a PCR amplification step. Sequencing was performed using Illumina NovaSeq SP, yielding 150 bp paired-end reads and an average of 7.42 Gbp per sample for MF and RO samples. To capture expected higher diversity, tertiary wastewater and BAC filtrate libraries were sequenced approximately four times more deeply than other samples, averaging 26.75 Gbp per sample (Table S1). All sequencing was conducted at the Vincent J. Coates Genomic Sequencing Laboratory (UC Berkeley).

2.5. Amplicon Sequencing Data Processing. Amplicon data analysis using DADA2[23] was performed as part of a larger project previously described.[11] Briefly, forward reads were filtered and trimmed to remove PhiX, truncate to 251 nt, trim the first 5 nt, remove reads shorter than 200 bp, truncate sequences with quality score $\leq$ 10, and remove reads with expected errors > 1. Next, reads were denoised to produce amplicon sequence variants (ASVs), chimeras were removed, and reads were classified against SILVA (v132). Samples with fewer than 7500 reads were removed, resulting in the loss of one tertiary wastewater sample and two negative controls. Seven positive extraction controls were examined. Of these, six appeared as expected, while one low-biomass positive control (2500 total cells input) showed contamination with ASVs classified as *Lysobacter* sp. present in both negative controls (field, extraction, and library preparation blanks) and samples were investigated using DESeq2 with absolute read-count data. ASVs that were significantly enriched in negative controls over samples were removed from samples (n = 151), including the same ASV classified in the positive control as belonging to *Lysobacter* sp. This method prevented removal of important ASVs that may have cross-contaminated from samples into controls, while removing ASVs present in all samples (including in controls) that were likely due to reagent contamination. Subsequently, one tertiary wastewater sample with fewer than 300 reads remaining was dropped from the data set (Table S1).

After processing and decontamination, data was further analyzed in R (v.4.2.0). Rarefaction curves were used to confirm adequate sequencing depth using *vegan* (v2.6.2) (Figure S3). Three samples were removed from further amplicon sequencing analysis (Figure S4): one RO permeate bulk water sample from day 26 showed evidence of cross-contamination with membrane biofilm samples during amplicon library preparation but not metagenomic library preparation (Figure S5); the RO membrane biofilm permeate sample from day 328 produced extremely low biomass (Table S1), prompting both the feed and permeate samples from that day to be removed from both amplicon and metagenomic analyses.

Diversity was investigated with *phyloseq* (v.1.40.0)[24] using the Shannon index, as this diversity index is relatively more robust for processing data with removal of singletons.[25] For β-diversity analysis, read counts were normalized with three methods, yielding nearly identical results: percent relative

https://doi.org/10.1021/acsestengg.4c00665
ACS EST Engg. 2025, 5, 772−781

abundance was calculated by dividing by total reads per sample, rarefaction was applied using the rarefy_even_depth function in *phyloseq*, and CLR-transformation was carried out using the *microbiome* package (v1.18.0).[26] Bray−Curtis dissimilarity was calculated for all 3 normalizations and PERMANOVA analysis was conducted on bulk water samples using vegan (adonis) to test for significant differences in community associated with sampling day and location. β-Diversity was visualized with nonmetric multidimensional scaling (NMDS) and principal coordinate analysis of Bray−Curtis dissimilarity calculated on rarefied data. Results of β-diversity analyses were nearly identical using percent-normalized and Aitchison distance-based methods (not shown).

The recurrent ASVs present after each treatment process were defined as those present at ≥0.05% relative abundance in at least two samples from that process, and intersections were visualized using *VennDiagram* (v.1.7.3) in R.

2.6. Metagenomic Data Analysis. Illumina adapters and PhiX were removed from raw reads using bbmap v38.79 (sourceforge.net/projects/bbmap/) and sequences were quality trimmed using sickle (https://github.com/najoshi/sickle). Quality control results were visually inspected using fastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads from each sample were individually assembled using IDBA-UD with a precorrection step.[27] Taxonomic analysis of contigs was performed with Kaiju[28] using the nr_euk database version r2021-02-24. Within the Anvi'o v7.1 environment,[29] Prodigal[30] was used to call open reading frames on scaffolds greater than 1 kbp, and single copy genes and rRNAs were identified using hmmscan.[31]

To assess and remove probable contaminants from metagenomic assemblies, reads from the field negative controls, extraction blanks, and the lowest-biomass positive control (2500 total cells input) were mapped to each assembly. For each assembly, Anvi'o was used to profile these mappings alongside the self-mapping of the reads used in the assembly. Contaminant scaffolds were identified with a custom script (find_contams.py) according to coverage, standard deviation of coverage, and detection in samples and controls (similar to Dai et al.[32]). The resulting collection of suspected contaminant scaffolds was manually reviewed using anvi-refine, and contaminant scaffolds were removed from further analysis by splitting the noncontaminants into new Anvi'o databases. Notably, two recurring contaminants were identified by kaiju as *Lysobacter enzymogenes* M497−1 and *Asaccharospora irregularis*, and these findings were consistent with the 16S decontamination results. This species of *Lysobacter* is used industrially to produce lysozyme for DNA extraction, which was included in this study (see Section 2.3) and was likely the source of this contamination.[33]

Pairwise MASH distances between read sets were calculated for all samples using mash v2.2.[34] These distances between samples confirmed that cross-mapping between samples from the same treatment process would likely assist with binning (Figure S5). Binning was performed for RO permeate, stored MF/UF, and RO membrane samples using DASTool[35] followed by dRep[36] and is further described in the Supporting Methods. Index of replication scores (iRep values) for each MAG and sample were calculated by inStrain[37,38] and subsequently filtered to include only those meeting the following conditions: non-normalized coverage depth was ≥5× in the sample, coverage breadth was >90% in the sample, the MAG had 175 scaffolds or less per Mbp of sequence, and completeness and contamination were estimated by CheckM as >75 and <5%, respectively. Functional annotation was performed using hmmscan with KEGG KOFAM profiles in Anvi'o (KOFAM HMMs from December 23, 2020) and investigated using anvi-estimate-metabolism. KEGG modules differentially enriched between MF and RO samples were determined with anvi-compute-functional-enrichment. MAG-specific in depth metabolic analysis used the KOFAM-based annotation table generated by anvi-summarize (Table S4) and the interactive KEGG map outputs from the web server version of KofamKOALA (https://www.genome.jp/tools/kofamkoala/).[39]

2.8. Data and Code Availability. Raw amplicon sequence data and metagenomic data were deposited at NCBI under project number PRJNA896721 and biosample numbers for samples included in this study are shown in Table S1. Code, including decontamination workflows, is available at https://github.com/rosekantor/awtp2 and https://github.com/rosekantor/metagenomics-tools.

3. RESULTS AND DISCUSSION

High-throughput sequencing was performed on bulk water samples collected from sites across the treatment train over three months (Figure 1) and on biofilm samples from one RO membrane harvested after the bulk water sampling period. Sequencing, quality control, and decontamination yielded 24 amplicon sequence data sets containing 3597 amplicon sequence variants (ASVs), 17 metagenomes (Table S1), and, for MF and RO permeates only, 117 metagenome-assembled genomes (MAGs) (Table S2). We previously reported total and intact flow cytometric cell counts for effluent from each treatment process.[9] Using these data, we calculated the percentage of intact cells to provide insight into the degree to which the sequence data represent potentially viable cells (Figures 2A and S1, Table S1). The fraction of potentially viable cells was similar for tertiary treated wastewater, ozone-BAC effluent, and MF effluent (48−100%), but after RO, a significant drop was observed (32−85% viable; $p = 0.012$ for pairwise t test with Benjamini Hochberg adjustment, between MF and RO), and after AOP, most cells were not viable (0−19%). Hence, AOP effluent samples were not collected for sequencing.

3.1. RO Microbial Community in the Context of the Full Treatment Train. Shannon's index of α diversity increased between tertiary treated wastewater and ozone-BAC effluent but was substantially lower in the MF/UF storage tank (here labeled as "MF") and RO permeate (Figure 2A). This trend in diversity was paralleled by the drop in cell counts and nutrient availability,[9] consistent with findings from another advanced treatment facility.[10] Although α diversities of MF and RO permeate samples were not significantly different from one another (Figure S6), pronounced differences in the microbial compositions were observed between each treatment step (PERMANOVA comparing bulk water samples only: $R\wedge2 = 0.5$, $p = 0.001$; Figures 2B and S9). Notably, this result was observed despite the use of different extraction kits for two of the RO samples and one of the BAC samples. Group dispersions, representing the degree of within-site similarity, were significantly different between treatment steps ($p <$ 9.999e-05), which may have contributed to the PERMANOVA result. In particular, the microbial composition of the water changed much more in the tertiary wastewater and ozone-BAC

https://doi.org/10.1021/acsestengg.4c00665
ACS EST Engg. 2025, 5, 772−781

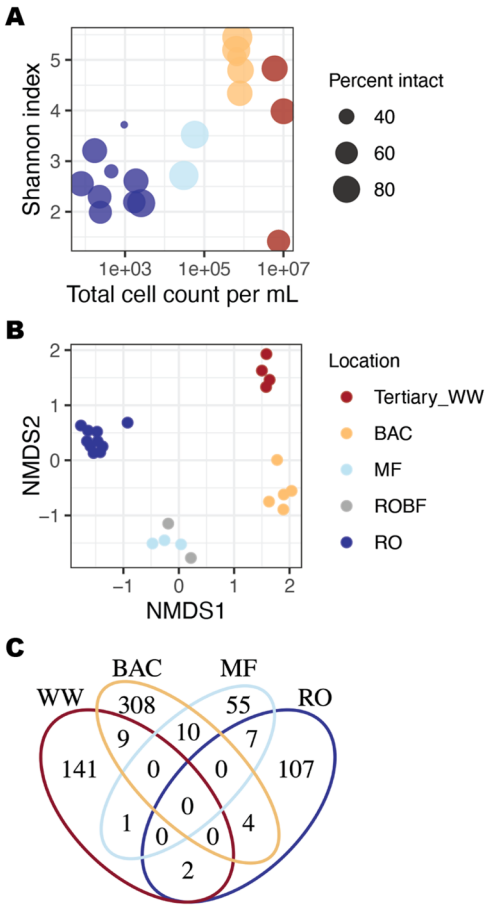


Figure 2. (A) Total cell count per mL and number of observed amplicon sequence variants (ASVs) decreased after each treatment step (color). Percent of total cells that were intact (point size) remained similar for all treatment processes except RO (see Figure S1). Note that sample volumes differed across the treatment train (see Table S1), and no samples of tertiary wastewater or MF permeate were collected for flow cytometry on Oct 10. (B) Nonmetric multidimensional scaling ordination of ASV abundances across samples using Bray–Curtis dissimilarity of rarefied data (stress = 0.11). This plot includes all bulk water samples used in PERMANOVA analysis as well as biomass samples from the RO membrane feed and permeate sides ("ROBF"). (C) Venn diagram of recurrent ASVs by sample type, defined as presence in at least two bulk water samples of the same type at >0.05% relative abundance.

effluent than in the MF and RO permeates over the three-month sampling period (Figure S10). The RO permeate microbial composition was consistent over time and also between two parallel RO trains with 2 and 3 stages (Figures 3 and S11), which suggests that the RO produced consistent microbial community regardless of temporal effects on wastewater quality or number of stages.

3.2. Microbial Identity and Origins of Microorganisms in the Post-RO Bulk Water. The phylum-level taxonomic composition of microorganisms identified across the treatment train shifted from Proteobacteria and Patescibacteria to Proteobacteria and Actinobacteria (Figure S7). At the family level, Burkholderiaceae and Mycobacteriaceae species were at high relative abundance in the RO permeate (Figures 3 and S8). The two most abundant ASVs were classified as *Aquabacterium* sp. (family Burkholderiaceae) and *Mycobacterium* sp. (family Mycobacteriaceae), and accounted for an average of 29 and 14% of the amplicon reads, respectively (Figure S11).

To understand whether the upstream treatment processes were the source of microorganisms in the RO permeate, we filtered the data set to include high-quality observations of ASVs or "recurrent ASVs" (see Section 2.5). This was done to avoid the possibility of misinterpreting one-off instances of low-abundance cross-contamination as recurrences of an ASV across treatment processes. The microbial composition at each sampling point was almost entirely distinct from that of other sampling points (Figure 2C). Recurrent ASVs were rarely shared between sequential processes (tertiary WW and BAC, BAC and MF, MF, and RO), and even more rarely shared between nonsequential processes. These results support that each unit process generally acted as a barrier to and selection pressure for a unique microbial community. No recurrent ASVs were shared between more than two processes, suggesting there was little to no consistent carry-over of bacteria across the treatment train. Specifically, only 2 recurrent ASVs were overlapping between tertiary wastewater and RO although 11 nonrecurrent ASVs were observed in both sample types (Figure S11). The presence of these ASVs in both samples may be due to (1) true treatment breakthrough, (2) cross-contamination during amplicon library preparation (see Section 3.4), or (3) the short size of the 16S V4 amplicon (which may be insufficient to distinguish closely related but nonidentical bacteria). Notably, of the two most abundant ASVs in RO permeate, the *Aquabacterium*- classified ASV was not detected in any upstream samples, but the *Mycobacterium*-classified ASV was seen in tertiary wastewater and BAC filtrate at low relative abundance. Similar to the amplicon sequencing results, no MAGs from MF and RO were detected in upstream tertiary wastewater samples by read-mapping when a coverage breadth threshold of 10% was applied (Figure S12).

Another potential source of microorganisms in RO permeate is the RO membranes themselves. To assess this possibility, we analyzed biomass collected when an RO membrane was autopsied to examine membrane fouling (Figure S2 and "ROBF" in Figure 2B). At the time of the autopsy, 27 days after the final bulk water sampling event, the RO train showed a 42.2% drop in flow (Elise Chen, Trussell Technologies, personal communication). This indicates extensive fouling, likely due to the lack of chloramine residual in the MF and RO feedwaters. Biomass collected from the retentate (feed) and permeate sides of the RO membrane shared most ASVs and MAGs (Figure S13 and Table S2), although the extracted DNA concentrations differed by over 2 orders of magnitude (Table S1). The shared microorganisms may be due to membrane breaches, but we cannot rule out the possibility of cross-contamination during the membrane autopsy and sample collection process. Overall, the RO membrane samples had similar profiles to the MF bulk water (stored MF and UF) just upstream of the RO membrane (Figure 2B) and shared few ASVs with the RO permeate samples (Figure S13). The abundant *Mycobacterium*-classified ASV was found in the RO permeate biofilm but not in the retentate biofilm. Of the abundant ASVs (>0.5% relative abundance) found in RO permeate bulk water, only one was detected on both feed and permeate sides of the RO membrane (Figure S11).

Although there are a small number of overlaps between upstream microorganisms and RO permeate, we speculate that the majority of microorganisms in the RO permeate were not likely derived from active membrane breakthrough (i.e.,

https://doi.org/10.1021/acsestengg.4c00665
ACS EST Engg. 2025, 5, 772–781

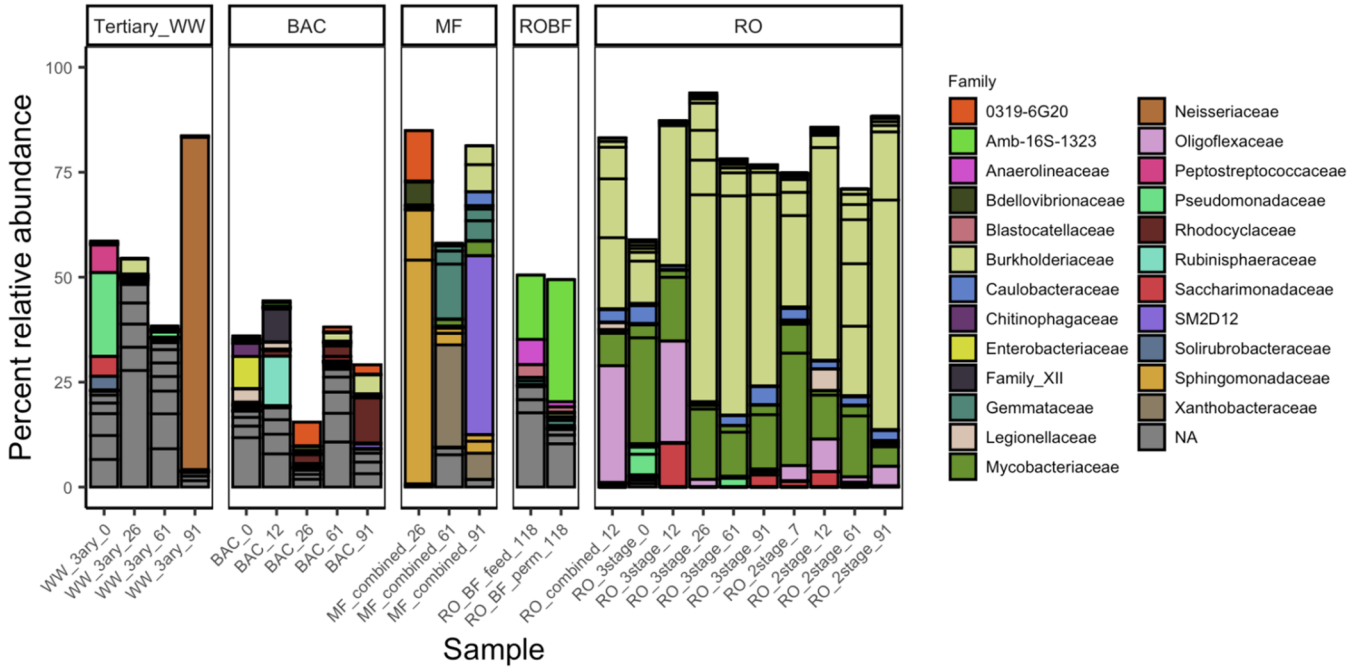


Figure 3. Family level taxonomy of most prevalent ASVs changed across the treatment train. Each segment of a bar represents a single ASV, colored by Family. ASVs are included if present in at least one sample at >3% relative abundance, so bars do not necessarily display all ASVs in each sample. Sample names on *x*-axis indicate the location and day of sampling, where day 0 is the date of the first sample taken in the study. Samples are grouped by treatment process; RO membrane biofilm ("ROBF") feed and permeate ("perm") refer to samples taken during a membrane autopsy.

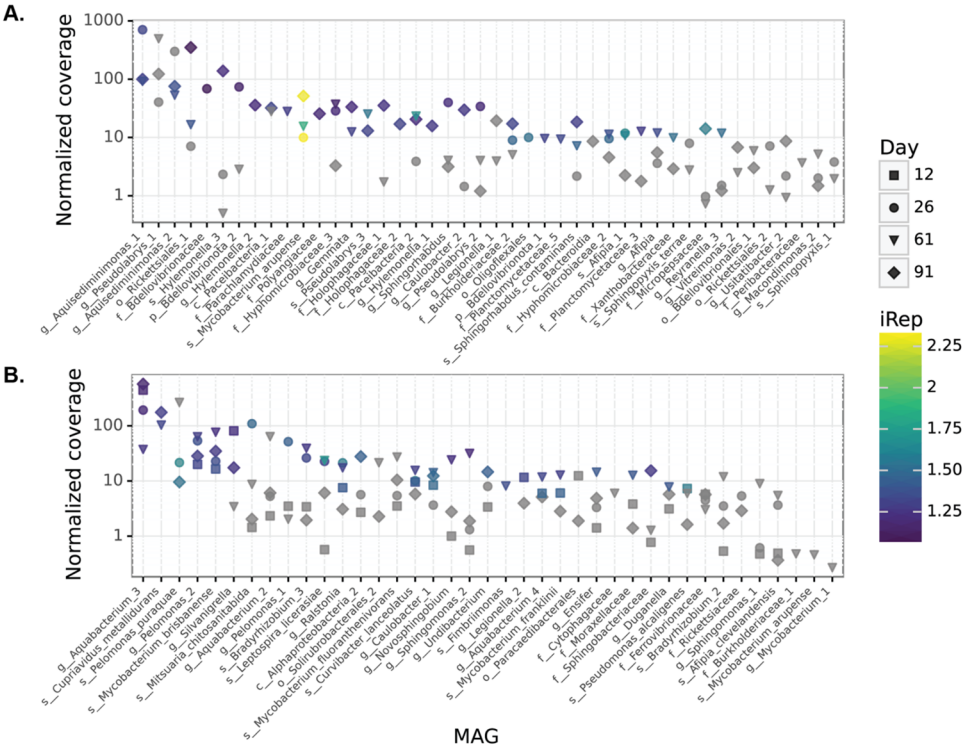


Figure 4. Relative abundances of MAGs shifted over time and were not always related to predicted growth rates. For (A) MF permeate and (B) RO permeate, the subset of MAGs present at >2× and >0× normalized coverage, respectively, are shown on a log10 scale. MAG coverages were normalized to account for sequencing depth (see Section 2). Points are colored by index of replication (iRep), and sample day is indicated by shape. Gray points are shown when the iRep score could not be calculated. An iRep score of 2 implies that, on average, all cells are doubling in the population represented by a given MAG.

breakthrough occurring during sampling), but rather persisted in the post-RO infrastructure (e.g., storage tank and piping downstream of the RO unit process) after an initial inoculation event. This seeding event may have occurred as part of the assembly of the RO train, during a previous membrane breakthrough, or via a gap in an O-ring seal.[40] A previous study

https://doi.org/10.1021/acsestengg.4c00665
ACS EST Engg. 2025, 5, 772−781

also reported the predominance of Burkholderiaceae after RO, potentially suggesting consistent selection for this group in RO permeate water.[41] The consistent abundance of the *Aquabacterium*-classified ASV and MAG suggests that a biofilm may have provided biomass to the bulk water during our sampling events, and other studies have also identified *Aquabacterium* sp. from drinking water biofilms.[42,43] Finally, other factors that may have influenced the RO permeate community include total facility operation time, membrane age, and the frequency of operational procedures such as membrane clean-in-place events, and their potential effects should be addressed by future research.

3.3. Survival and Growth of Microorganisms in the Oligotrophic RO Permeate Environment. An open question is whether bacteria in the RO permeate were actively growing, and if so, what metabolic strategies may be employed. Although subsequently eliminated by AOP, any bacterial growth in RO may be representative of the potential for downstream growth as water from the advanced purification facility is stored and further managed. To address this question, we previously analyzed the ratios of low-nucleic acid (LNA) to high-nucleic acid (HNA) cells in bulk water collected across the treatment train, revealing a higher proportion of (HNA) bacteria in the MF/UF and RO permeates compared to the BAC effluent.[9] We concluded that recent growth may have occurred in the MF/UF storage tank ("MF" in this study), as nucleic acid content per cell is sometimes an indicator of growth. Analysis of the microbial community offers an alternative possibility: there were more LNA bacteria in the BAC effluent than in MF and RO. These bacteria corresponded to small cell-size, small-genome symbionts (e.g., Patescibacteria from the Candidate Phyla Radiation; Figure S7)[44] that were present due to the abundance of host bacteria. In contrast, the oligotrophic environments of the MF and RO permeates likely did not provide enough biomass to support this additional layer of the microbial food web. Instead, we observed the prevalence of large-genome bacteria, particularly in the RO permeate (Table S2).

As another metric of growth, we used inStrain[38] to calculate index of replication scores (iRep).[37] These scores indicate the fold-rate of replication of populations represented by each MAG in each sample, and nearly all were below 2.0 in MF and RO samples (Figure 4). In MF permeate, the MAG classified as *Mycobacterium arupense* had one of the highest iRep scores, and, concordantly, this MAG increased in relative abundance over time (Figure 4A). Interestingly, the MAG with highest relative abundance in MF (*Aquisediminimonas* sp.) did not have a high iRep score at any time point, and decreased in relative abundance over the three sample dates. In RO permeate samples, the highly abundant Aquabacterium_3 MAG showed a low predicted growth rate (Figure 4B). Overall, the low iRep scores suggest low growth rates, which is aligned with the oligotrophic conditions of the membrane permeate. It is also possible that some of these bacteria were derived from slower growing biofilm on pipe walls downstream of the RO unit process,[45] as suggested above.

Next, we analyzed the predicted metabolisms of MAGs recovered from the MF and RO samples to explore potential growth strategies for these MAGs. Low molecular weight, neutral molecules, such as formaldehyde, are poorly removed by RO membranes,[46] but no MAGs encoded complete pathways for formaldehyde assimilation. Many MAGs contained amino acid and nucleic acid degradation pathways (Figure S12), suggesting possible growth via biomass recycling. A metabolic enrichment analysis showed that several nucleic acid degradation pathways were significantly enriched in RO MAGs relative to MF MAGs (Table S3), and smaller subset of MAGs encoded nitrogen fixation and ribulose-bisphosphate carboxylase-oxygenase (RuBisCO) suggesting the potential for CO_2 fixation (Figure S12). Of these, we focused on the Aquabacterium_3 MAG for in-depth metabolic prediction (Table S4), because it was prevalent in all RO permeate samples.

In agreement with previous culture-based work on *Aquabacterium* spp.,[43] the Aquabacterium_3 MAG encoded no pathways for starch or sugar utilization except for mannose, glucose, and fructose, none of which would be expected to be present in RO permeate. However, the MAG contained the complete pathway for β-oxidation of fatty acids, and cultured isolates have been reported to use a variety of small organic molecules as a carbon source. Further, this MAG encoded polyhydroxybutyrate biosynthesis, which could allow it to store carbon as inclusion bodies like those previously observed in isolates.[43] Additionally, we found a complete set of enzymes for carbon fixation via the Calvin cycle but no genes suggesting chemotrophic and phototrophic metabolism. As observed in cultured isolates, the MAG corresponded to a facultative aerobe, encoding oxidative phosphorylation, denitrification to nitrous oxide, and dissimilatory nitrate reduction to ammonia. Although not previously reported for *Aquabacterium* sp., this MAG encoded the capacity for nitrogen fixation, a trait that may be important if the RO permeate is nitrogen-limited. Finally, the MAG harbored genes for type IV pili and flagella, important for surface attachment and biofilm formation. Overall, this genome appears well-adapted for survival in RO permeate, and may have persisted in biofilm downstream from the RO membranes.

3.4. Study Limitations and Future Directions. Two important limitations to this work include the study site and the low biomass recovered in RO samples. First, one RO system is not likely to be representative of all, especially as upstream treatment processes and operating parameters will differ among (and potentially within) facilities. For example, during our sampling period, the facility was operated without a chloramine residual, which may have affected cell growth, nutrient availability, and the microbial communities observed. Second, the sequencing data was affected by reagent-associated contamination, as is common in low biomass studies.[47,48] These sequences were readily identified and removed by comparison to negative controls, and the extent of contamination in the metagenomic data was clearly related to the concentration of DNA resulting from extraction (Figure S14). Trends in contamination of the 16S rRNA gene amplicon sequencing data were less apparent due to the intermediate steps of PCR, library normalization, and pooling, but the same reagent-associated contaminants were identified during data decontamination (see Section 2). Owing to the plate-based PCR steps during amplicon library preparation, the amplicon data may also have been affected by cross-library contamination, which is much more difficult to account for than reagent-associated contamination.[49] Nevertheless, Miller et al.[9] previously analyzed our samples with qPCR, showing that DNA concentrations and 16S rRNA gene copy numbers were significantly higher in these RO samples compared to negative controls. Given the application of negative controls, careful

https://doi.org/10.1021/acsestengg.4c00665
ACS EST Engg. 2025, 5, 772−781

decontamination with manual review, and minimal overlap with other sample types (Figure 2C), we are reasonably confident that the RO permeate bulk water sequence data was representative of the samples collected.

4. CONCLUSIONS

A key contribution of this study is the description of RO permeate microorganisms, consistency of the microbial community, and a lack of potential sources for these organisms. While previous work demonstrated log removals of bacteria across the treatment train,[9] our microbial composition analysis shows that different organisms were present after each unit process. This implies that the majority of the microbial biomass observed was not "regrowth" but rather the takeover of a new set of organisms after each treatment process.[50] In this regard, it is perhaps more relevant to consider the removal of nutrients and changes to the nutrient profile across the treatment train. Those changes, while unmeasured in this study, appear to have affected the microbial profile through selection of specific groups. Although we were not able to sample after the AOP treatment or downstream steps that would be present in a full-scale system (e.g., remineralization, storage, or distribution), we suggest that the microorganisms observed in RO permeate samples may be representative of the type of community that could grow after treatment, given the nutrient profile of the water. Critically, the results of feeding RO-treated, highly purified water into a drinking water distribution system or storing this water are still largely unknown. These downstream phenomena are difficult to simulate using pipe loops[11] or bench-scale reactors, and likely will not be reasonably known until such water enters a full-scale system and can be sampled in situ.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsestengg.4c00665.

Providing data on flow cytometry, sampling, amplicon sequencing, and metagenomic sequencing; supporting methods detailing the metagenomic binning process Figures (S1−S14) (PDF)

Describing sampling, sequence data, MAG abundances, and predicted metabolisms (Tables S1−S4) (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

Rose S. Kantor − *Department of Civil and Environmental Engineering, University of California, Berkeley, California 94720, United States; National Science Foundation Engineering Research Center for Re-inventing the Nation's Urban Water Infrastructure (ReNUWIt), Berkeley, California 94720, United States;* Present Address: Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, California 94550, United States; orcid.org/0000-0002-5402-8979; Email: kantor4@llnl.gov

Kara L. Nelson − *Department of Civil and Environmental Engineering, University of California, Berkeley, California 94720, United States; National Science Foundation Engineering Research Center for Re-inventing the Nation's Urban Water Infrastructure (ReNUWIt), Berkeley, California 94720, United States;* orcid.org/0000-0001-8899-2662; Email: karanelson@berkeley.edu

Authors

Lauren C. Kennedy − *Department of Civil and Environmental Engineering, University of California, Berkeley, California 94720, United States; National Science Foundation Engineering Research Center for Re-inventing the Nation's Urban Water Infrastructure (ReNUWIt), Berkeley, California 94720, United States; Department of Civil Engineering, University of Texas at El Paso, El Paso, Texas 79968, United States;* orcid.org/0000-0002-4451-2361

Scott E. Miller − *Department of Civil and Environmental Engineering, University of California, Berkeley, California 94720, United States; National Science Foundation Engineering Research Center for Re-inventing the Nation's Urban Water Infrastructure (ReNUWIt), Berkeley, California 94720, United States;* orcid.org/0000-0003-0321-7806

Jorien Favere − *Center for Microbial Ecology and Technology (CMET), Ghent University, 9000 Gent, Belgium; Centre for Advanced Process Technology for Urban Resource Recovery (CAPTURE), Ghent University, 9000 Gent, Belgium*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsestengg.4c00665

Author Contributions

CRediT: **Rose S Kantor** conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing - original draft, writing - review & editing; **Lauren C. Kennedy** conceptualization, data curation, writing - review & editing; **Scott E. Miller** conceptualization, data curation, investigation, methodology; **Jorien Favere** data curation, formal analysis, visualization; **Kara L. Nelson** conceptualization, funding acquisition, supervision, writing - review & editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Trussell Technologies for assistance with sample collection, Dr. Rohan Sachdeva and the Banfield-IGI computing cluster for access to computational resources, and the UC Berkeley QB3 Functional Genomics Laboratory and Genomic Sequencing Laboratory for library preparation and sequencing. This work was supported by the National Science Foundation (NSF) through the Engineering Research Center for Reinventing the Nation's Urban Water Infrastructure (ReNUWIt) EEC-1028968 and Environmental Protection Agency Science to Achieve Results Graduate Fellowship (EPA STAR no. 91782901-0) to S.E. Miller.

■ REFERENCES

(1) Garrick, D.; Hall, J. W. Water Security and Society: Risks, Metrics, and Pathways. *Annu. Rev. Environ. Resour.* **2014**, *39*, 611−639.

(2) California's Water Supply Strategy; State of California 2022 https://resources.ca.gov/-/media/CNRA-Website/Files/Initiatives/Water-Resilience/CA-Water-Supply-Strategy.pdf. (accessed June 25, 2024).

(3) Du Pisani, P.; Menge, J. G. Direct Potable Reclamation in Windhoek: A Critical Review of the Design Philosophy of New

Goreangab Drinking Water Reclamation Plant. *Water Sci. Technol.: Water Supply* **2013**, *13* (2), 214−226.
(4) Maseeh, G. P.; Russell, C. G.; Villalobos, S. L.; Balliew, J. E.; Trejo, G. El Paso's Advanced Water Purification Facility: A New Direction in Potable Reuse. *J. AWWA* **2015**, *107* (11), 36−45.
(5) All Options on the Table: Lessons from the Journeys of Others; Water Services Association of Australia 2019 https://wsaa.asn.au/Web/Web/News-and-Resources/Reports/All-options-on-the-table--lessons-from-the-journeys-of-others.aspx.
(6) Jeffrey, P.; Yang, Z.; Judd, S. J. The Status of Potable Water Reuse Implementation. *Water Res.* **2022**, *214*, No. 118198.
(7) Marron, E. L.; Mitch, W. A.; von Gunten, U.; Sedlak, D. L. A Tale of Two Treatments: The Multiple Barrier Approach to Removing Chemical Contaminants During Potable Water Reuse. *Acc. Chem. Res.* **2019**, *52* (3), 615−622.
(8) Yasui, M.; Iso, H.; Torii, S.; Matsui, Y.; Katayama, H. Applicability of Pepper Mild Mottle Virus and Cucumber Green Mottle Mosaic Virus as Process Indicators of Enteric Virus Removal by Membrane Processes at a Potable Reuse Facility. *Water Res.* **2021**, *206*, No. 117735.
(9) Miller, S.; Greenwald, H.; Kennedy, L. C.; Kantor, R. S.; Jiang, R.; Pisarenko, A.; Chen, E.; Nelson, K. L. Microbial Water Quality through a Full-Scale Advanced Wastewater Treatment Demonstration Facility. *ACS EST Eng.* **2022**, *2*, 2206. acsestengg.2c00198
(10) Kantor, R. S.; Miller, S. E.; Nelson, K. L. The Water Microbiome Through a Pilot Scale Advanced Treatment Facility for Direct Potable Reuse. *Front. Microbiol.* **2019**, *10*, 993.
(11) Kennedy, L. C.; Miller, S. E.; Kantor, R. S.; Greenwald, H.; Adelman, M. J.; Seshan, H.; Russell, P.; Nelson, K. L. Stay in the Loop: Lessons Learned about the Microbial Water Quality in Pipe Loops Transitioned from Conventional to Direct Potable Reuse Water. *Environ. Sci.: Water Res. Technol.* **2023**, *9* (5), 1436−1454.
(12) Miller, S. E.; Rodriguez, R. A.; Nelson, K. L. Removal and Growth of Microorganisms across Treatment and Simulated Distribution at a Pilot-Scale Direct Potable Reuse Facility. *Environ. Sci.: Water Res. Technol.* **2020**, *6* (5), 1370−1387.
(13) Sousi, M.; Salinas-Rodriguez, S. G.; Liu, G.; Schippers, J. C.; Kennedy, M. D.; van der Meer, W. Measuring Bacterial Growth Potential of Ultra-Low Nutrient Drinking Water Produced by Reverse Osmosis: Effect of Sample Pre-Treatment and Bacterial Inoculum. *Front. Microbiol.* **2020**, *11*, 791.
(14) Jafari, M.; Vanoppen, M.; van Agtmaal, J. M. C.; Cornelissen, E. R.; Vrouwenvelder, J. S.; Verliefde, A.; van Loosdrecht, M. C. M.; Picioreanu, C. Cost of Fouling in Full-Scale Reverse Osmosis and Nanofiltration Installations in the Netherlands. *Desalination* **2021**, *500*, No. 114865.
(15) Garner, E.; Inyang, M.; Garvey, E.; Parks, J.; Glover, C.; Grimaldi, A.; Dickenson, E.; Sutherland, J.; Salveson, A.; Edwards, M. A.; Pruden, A. Impact of Blending for Direct Potable Reuse on Premise Plumbing Microbial Ecology and Regrowth of Opportunistic Pathogens and Antibiotic Resistant Bacteria. *Water Res.* **2019**, *151*, 75−86.
(16) Harb, M.; Wang, P.; Zarei-Baygi, A.; Plumlee, M. H.; Smith, A. L. Background Antibiotic Resistance and Microbial Communities Dominate Effects of Advanced Purified Water Recharge to an Urban Aquifer. *Environ. Sci. Technol. Lett.* **2019**, *6* (10), 578−584.
(17) Nocker, A.; Schulte-Illingheim, L.; Müller, H.; Rohn, A.; Zimmermann, B.; Gaba, A.; Nahrstedt, A.; Mohammadi, H.; Tiemann, Y.; Krömer, K. Microbiological Changes along a Modular Wastewater Reuse Treatment Process with a Special Focus on Bacterial Regrowth. *J. Water Reuse Desalin.* **2020**, *10* (4), 380−393.
(18) Stamps, B. W.; Leddy, M. B.; Plumlee, M. H.; Hasan, N. A.; Colwell, R. R.; Spear, J. R. Characterization of the Microbiome at the World's Largest Potable Water Reuse Facility. *Front. Microbiol.* **2018**, *9*, 2435.
(19) Stamps, B. W.; Spear, J. R. Identification of Metagenome-Assembled Genomes Containing Antimicrobial Resistance Genes, Isolated from an Advanced Water Treatment Facility. *Microbiol. Resour. Announce.* **2020**, *9* (14), 10−1128, DOI: 10.1128/MRA.00003-20.
(20) Trussell, R. S.; Pisarenko, A. N. Process Benefits of Ozone/BAC as Pretreatment to Membrane-Based Advanced Treatment for Direct Potable Reuse. *Water Reuse* **2024**, *14* (2), 146−159.
(21) Smith, C. M.; Hill, V. R. Dead-End Hollow-Fiber Ultrafiltration for Recovery of Diverse Microbes from Water. *Appl. Environ. Microbiol.* **2009**, *75* (16), 5284−5289.
(22) Kozich, J. J.; Westcott, S. L.; Baxter, N. T.; Highlander, S. K.; Schloss, P. D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl. Environ. Microbiol.* **2013**, *79* (17), 5112−5120.
(23) 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.
(24) McMurdie, P. J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **2013**, *8* (4), No. e61217.
(25) Chiarello, M.; McCauley, M.; Villéger, S.; Jackson, C. R. Ranking the Biases: The Choice of OTUs vs. ASVs in 16S rRNA Amplicon Data Analysis Has Stronger Effects on Diversity Measures than Rarefaction and OTU Identity Threshold. *PLoS One* **2022**, *17* (2), No. e0264443.
(26) Lahti, L.; Shetty, S. Microbiome R package. 2023 DOI: 10.18129/B9.bioc.microbiome.
(27) Peng, Y.; Leung, H. C. M.; Yiu, S. M.; Chin, F. Y. L. IDBA-UD: A de Novo Assembler for Single-Cell and Metagenomic Sequencing Data with Highly Uneven Depth. *Bioinformatics* **2012**, *28* (11), 1420−1428.
(28) Menzel, P.; Ng, K. L.; Krogh, A. Fast and Sensitive Taxonomic Classification for Metagenomics with Kaiju. *Nat. Commun.* **2016**, *7* (1), No. 11257.
(29) Eren, A. M.; Kiefl, E.; Shaiber, A.; Veseli, I.; Miller, S. E.; Schechter, M. S.; Fink, I.; Pan, J. N.; Yousef, M.; Fogarty, E. C.; Trigodet, F.; Watson, A. R.; Esen, Ö. C.; Moore, R. M.; Clayssen, Q.; Lee, M. D.; Kivenson, V.; Graham, E. D.; Merrill, B. D.; Karkman, A.; Blankenberg, D.; Eppley, J. M.; Sjödin, A.; Scott, J. J.; Vázquez-Campos, X.; McKay, L. J.; McDaniel, E. A.; Stevens, S. L. R.; Anderson, R. E.; Fuessel, J.; Fernandez-Guerra, A.; Maignien, L.; Delmont, T. O.; Willis, A. D. Community-Led, Integrated, Reproducible Multi-Omics with Anvi'o. *Nat. Microbiol.* **2021**, *6* (1), 3−6.
(30) Hyatt, D.; Chen, G.-L.; LoCascio, P. F.; Land, M. L.; Larimer, F. W.; Hauser, L. J. Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification. *BMC Bioinf.* **2010**, *11* (1), 119.
(31) Eddy, S. R. Accelerated Profile HMM Searches. *PLOS Comput. Biol.* **2011**, *7* (10), No. e1002195.
(32) Dai, Z.; Sevillano-Rivera, M. C.; Calus, S. T.; Bautista-de los Santos, Q. M.; Eren, A. M.; van der Wielen, P. W. J. J.; Ijaz, U. Z.; Pinto, A. J. Disinfection Exhibits Systematic Impacts on the Drinking Water Microbiome. *Microbiome* **2020**, *8* (1), 42.
(33) Takami, H.; Toyoda, A.; Uchiyama, I.; Itoh, T.; Takaki, Y.; Arai, W.; Nishi, S.; Kawai, M.; Shin-ya, K.; Ikeda, H. Complete Genome Sequence and Expression Profile of the Commercial Lytic Enzyme Producer Lysobacter Enzymogenes M497−1. *DNA Res.* **2017**, *24* (2), 169−177.
(34) Ondov, B. D.; Treangen, T. J.; Melsted, P.; Mallonee, A. B.; Bergman, N. H.; Koren, S.; Phillippy, A. M. Mash: Fast Genome and Metagenome Distance Estimation Using MinHash. *Genome Biol.* **2016**, *17* (1), 132.
(35) 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.
(36) 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.
(37) Brown, C. T.; Olm, M. R.; Thomas, B. C.; Banfield, J. F. Measurement of Bacterial Replication Rates in Microbial Communities. *Nat. Biotechnol.* **2016**, *34* (12), 1256−1263.
(38) Olm, M. R.; Crits-Christoph, A.; Bouma-Gregson, K.; Firek, B. A.; Morowitz, M. J.; Banfield, J. F. inStrain Profiles Population Microdiversity from Metagenomic Data and Sensitively Detects Shared Microbial Strains. *Nat. Biotechnol.* **2021**, *39* (6), 727−736.
(39) Aramaki, T.; Blanc-Mathieu, R.; Endo, H.; Ohkubo, K.; Kanehisa, M.; Goto, S.; Ogata, H. KofamKOALA: KEGG Ortholog Assignment Based on Profile HMM and Adaptive Score Threshold. *Bioinformatics* **2020**, *36* (7), 2251−2252.
(40) Fujioka, T.; Hoang, A. T.; Ueyama, T.; Nghiem, L. D. Integrity of Reverse Osmosis Membrane for Removing Bacteria: New Insight into Bacterial Passage. *Environ. Sci. Water Res. Technol.* **2019**, *5* (2), 239−245.
(41) Fujioka, T.; Boivin, S. Assessing Bacterial Infiltration through Reverse Osmosis Membrane. *Environ. Technol. Innovation* **2020**, *19*, No. 100818.
(42) Kalmbach, S.; Manz, W.; Bendinger, B.; Szewzyk, U. *In Situ* Probing Reveals *Aquabacterium Commune* as a Widespread and Highly Abundant Bacterial Species in Drinking Water Biofilms. *Water Res.* **2000**, *34* (2), 575−581.
(43) Kalmbach, S.; Manz, W.; Wecke, J.; Szewzyk, U. Aquabacterium Gen. Nov., with Description of Aquabacterium Citratiphilum Sp. Nov., Aquabacterium Parvum Sp. Nov. and Aquabacterium Commune Sp. Nov., Three in Situ Dominant Bacterial Species from the Berlin Drinking Water System. *Int. J. Syst. Evol. Microbiol.* **1999**, *49* (2), 769−777.
(44) Proctor, C. R.; Besmer, M. D.; Langenegger, T.; Beck, K.; Walser, J.-C.; Ackermann, M.; Bürgmann, H.; Hammes, F. Phylogenetic Clustering of Small Low Nucleic Acid-Content Bacteria across Diverse Freshwater Ecosystems. *ISME J.* **2018**, *12* (5), 1344−1359.
(45) Boe-Hansen, R.; Albrechtsen, H.-J.; Arvin, E.; Jørgensen, C. Bulk Water Phase and Biofilm Growth in Drinking Water at Low Nutrient Conditions. *Water Res.* **2002**, *36* (18), 4477−4486.
(46) Breitner, L. N.; Howe, K. J.; Minakata, D. Effect of Functional Chemistry on the Rejection of Low-Molecular Weight Neutral Organics through Reverse Osmosis Membranes for Potable Reuse. *Environ. Sci. Technol.* **2019**, *53* (19), 11401−11409.
(47) Karstens, L.; Asquith, M.; Davin, S.; Fair, D.; Gregory, W. T.; Wolfe, A. J.; Braun, J.; McWeeney, S. Controlling for Contaminants in Low-Biomass 16S rRNA Gene Sequencing Experiments. *MSystems* **2019**, *4* (4), 10−1128, DOI: 10.1128/mSystems.00290-19.
(48) Salter, S. J.; Cox, M. J.; Turek, E. M.; Calus, S. T.; Cookson, W. O.; Moffatt, M. F.; Turner, P.; Parkhill, J.; Loman, N. J.; Walker, A. W. Reagent and Laboratory Contamination Can Critically Impact Sequence-Based Microbiome Analyses. *BMC Biol.* **2014**, *12* (1), 87.
(49) Lou, Y. C.; Hoff, J.; Olm, M. R.; West-Roberts, J.; Diamond, S.; Firek, B. A.; Morowitz, M. J.; Banfield, J. F. Using Strain-Resolved Analysis to Identify Contamination in Metagenomics Data. *Microbiome* **2023**, *11* (1), 36.
(50) Blair, M. F.; Garner, E.; Ji, P.; Pruden, A. What Is the Difference between Conventional Drinking Water, Potable Reuse Water, and Nonpotable Reuse Water? A Microbiome Perspective. *Environ. Sci. Technol.* **2024**, *58* (38), 16877−16890.

https://doi.org/10.1021/acsestengg.4c00665
ACS EST Engg. 2025, 5, 772−781