Source-to-Tap Investigation of the Occurrence of Nontuberculous Mycobacteria in a Full-Scale Chloraminated Drinking Water System

Authors

Katherine S. Dowdell,[1] Sarah C. Potgieter,[1] Kirk Olsen,[1] Soojung Lee,[1] Matthew Vedrin,[1] Lindsay J. Caverly,[2] John J. LiPuma,[2] and Lutgarde Raskin[*,1]

Author Affiliations

[1]Department of Civil and Environmental Engineering, University of Michigan, 1353 Beal Ave., Ann Arbor, MI 48109, USA

[2]Department of Pediatrics, University of Michigan Medical School, 8323 MSRB III, SPC5646, 1150 W. Med Cntr Dr., Ann Arbor, MI 48109, USA

[*]**Corresponding Author**

Lutgarde Raskin (raskin@umich.edu)

Keywords

Nontuberculous mycobacteria, drinking water, opportunistic pathogens, *Mycobacterium*, building plumbing

Abstract

Nontuberculous mycobacteria (NTM) in drinking water are a significant public health concern. However, an incomplete understanding of the factors that influence the occurrence of NTM in drinking water limits our ability to characterize risk and prevent infection. This study sought to evaluate the influence of season and water treatment, distribution, and stagnation on NTM in drinking water. Samples were collected source-to-tap in a full-scale, chloraminated drinking water system approximately monthly from December 2019 to November 2020. NTM were characterized using culture-dependent (plate culture with matrix-assisted laser desorption ionization-time of flight mass spectrometry [MALDI-TOF MS] isolate analysis) and culture-independent methods (quantitative PCR and genome resolved metagenomics). Sampling locations included source waters, three locations within the treatment plant, and five buildings receiving water from the distribution system. Building plumbing samples consisted of first draw, five-minute flush, and full flush cold-water samples. As the study took place during the COVID-19 pandemic, the influence of reduced water usage in three of the five buildings was also investigated. The highest concentrations of NTM source-to-tap were found in the summer first draw building water samples (10^7 gene copies/L), which also had the lowest chloramine concentrations. Flushing was found to be effective for reducing NTM and restoring disinfectant residuals, though flush times necessary to improve water quality varied by building. Clinically-relevant NTM species, including *Mycobacterium avium*, were recovered via plate culture, with increased occurrence observed in buildings with higher water age. Four of five NTM metagenome-assembled genomes were identified to the species level and matched identified isolates.

Importance

NTM infections are increasing in prevalence, difficult to treat, and associated with high mortality rates. Our lack of understanding of the factors that influence NTM occurrence in drinking water limits our ability to prevent infections, accurately characterize risk, and focus remediation efforts. In this study, we comprehensively evaluated NTM in a full-scale drinking water system, showing that various steps in treatment and distribution influence NTM presence. Stagnant building water contained the highest NTM concentrations source-to-tap and was associated with low disinfectant residuals. We illustrated the differences in NTM detection and characterization obtained from culture-based and culture-independent methods, highlighting the complementarity between these

52 approaches. We demonstrated that focusing NTM mitigation efforts in building plumbing systems,
53 which have the highest NTM concentrations source-to-tap, has potential for immediate positive
54 effects. We also identified steps during treatment that increase NTM levels, which provides
55 beneficial information for utilities seeking to reduce NTM in finished water.
56

Introduction

Nontuberculous mycobacteria (NTM) respiratory infections are a significant and growing health concern, with numerous recent studies reporting increasing prevalence (1–3). The impact of NTM infections is illustrated in high associated mortality rates and healthcare costs compared to infections by enteric waterborne pathogens (4–6). Though the sources of NTM infections are often difficult to determine, drinking water exposure is believed to be a major route of infection (7, 8). However, the factors that shape NTM concentrations at the tap are poorly understood. This limitation is due to several reasons, including the complexity of quantification methods, the lack of regulations governing NTM in drinking water, and the difficulty in linking specific environmental sources to infections (9). A greater understanding of NTM occurrence, species composition of NTM populations, and factors that contribute to higher NTM levels in drinking water is crucial for the development of quantitative risk assessments and risk mitigation strategies (10, 11).

Building plumbing characteristics have been shown to greatly influence tap water quality. Building plumbing properties, including high surface-to-volume ratios, periods of stagnation, and equilibration with building temperatures, create an environment that contributes to the growth of biofilms and high concentrations of certain opportunistic human pathogens (OPs), including NTM (12, 13). Although previous studies have found that building plumbing may contribute to high concentrations of NTM, few studies have evaluated the impacts of flushing, seasonal variations, and water quality parameters (14–17). Previous studies have found concentrations of NTM in building plumbing as high as 10^5 - 10^7 gene copies per liter (15, 18–20). The distribution system has also been shown to influence NTM, with higher concentrations of NTM or greater occurrence of clinically-relevant species linked to higher distribution system residence times (14, 21). Source water type (22) and drinking water treatment processes may also influence the NTM entering the distribution system, though few studies have investigated these topics in full-scale systems (23–27). Another factor linked to NTM occurrence is the type of secondary disinfectant used, with several studies reporting increased rates of NTM detection or higher NTM concentrations in distribution systems using monochloramine (22, 28–30). The COVID-19 pandemic brought additional concerns about OPs in building water due to low water use (31). Although most studies

87 focused only on *Legionella pneumophila* (32–34), one study reported *Mycobacterium avium*
88 complex up to 10^5 gene copies per liter in stagnant building plumbing water (16).
89
90 Culture-based methods have long been the primary means for detecting NTM in environmental
91 samples. Culture-based methods are often favored because they detect viable bacteria that can be
92 used for additional analyses. However, existing culture-based methods for recovery of NTM from
93 drinking water lack standardization, are laborious, and, despite the use of decontamination or
94 antibiotic-containing media, are often limited in their specificity for NTM, meaning that additional
95 methods are required to confirm species identity (35–37). Further, species- and strain-level
96 variation in susceptibility to decontamination treatments and antibiotics may introduce bias in
97 quantifying NTM populations in mixed communities in environmental samples (36–38).
98 Quantitative PCR (qPCR) is often favored by researchers for quantifying NTM because it is rapid
99 and typically highly specific at the genus or species level (39–43). However, unless viability pre-
100 treatments are used (e.g. propidium monoazide or ethidium monoazide) (44–46), qPCR does not
101 provide information regarding viability. Recently, amplicon sequencing and shotgun
102 metagenomics have been used to determine the relative abundance of NTM in microbial
103 communities, with some studies reporting the enrichment of NTM in building plumbing biofilms
104 (14, 47–49). However, DNA sequencing methods may fail to detect low abundance
105 microorganisms like NTM, they do not show viability, and the PCR employed in amplicon
106 sequencing introduces bias (50, 51). Additionally, sequencing methods can only yield relative,
107 rather than absolute abundances unless quantitative metagenomics is employed (52, 53). NTM are
108 also difficult to lyse, which can impact DNA recovery and NTM detection (54). As methods for
109 NTM quantification are not standardized, cross-study comparisons are challenging, particularly
110 between studies that use culture-based methods versus those that use culture-independent methods.
111
112 The risk to human health associated with NTM in drinking water is also dependent on the species
113 present. Species within *Mycobacterium* range from non-pathogenic, to OPs (e.g., *M. avium* and
114 *Mycobacterium abscessus*), to non-environmental pathogens (e.g., *Mycobacterium tuberculosis*).
115 Although the risk of infection due to OPs is highly dependent on the susceptibility of the host, it
116 is estimated that approximately one dozen of the nearly 200 species of NTM are capable of causing
117 human infection (55). Additionally, regional differences have been observed regarding which

NTM species are most likely to cause human infection. Such regional differences could result in underestimating risk if the most appropriate species are not targeted (56, 57). Due to the difficulties in identifying NTM to the species- or complex-level, drinking water studies often quantify total NTM, with many targeting the ATP synthase subunit c (*atpE*) gene (15, 58–60), or only investigating one or a few species of concern, such as *M. avium* or *Mycobacterium intracellulare* (19, 22, 61, 62). Identification of NTM to the species level in colonies from plate culture is typically done using PCR and Sanger sequencing targeting genes such as β-subunit of RNA polymerase (*rpoB)* or heat shock protein 65 (*hsp65*) (63–66). Another alternative for colony identification is matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS), which is a method that analyzes proteins and generates spectra that are matched to a spectral database (67, 68). Advantages of MALDI-TOF MS include that it is rapid, accurate, and commonly used in clinical laboratories (69, 70). Culture-independent methods for species-level identification of NTM typically employ short-read or long-read amplicon sequencing, with targets such as the *rpoB* or *hsp65* gene (14, 47, 71).

Although previous studies have quantified NTM at various locations in drinking water systems (14, 15, 28, 30, 72), few have conducted full, source-to-tap assessments (73). This study sought to characterize the factors that shape NTM concentrations at various stages of treatment and distribution, including impacts of source water selection, individual treatment processes, distribution, and stagnation in building plumbing. To this end, samples were collected approximately monthly over one year in a full-scale chloraminated system. Samples were collected from source waters (river and well water), through the drinking water treatment plant (WTP), and in five buildings (Sites A-E) served by the WTP. The estimated distribution system water ages for the five buildings ranged from approximately 16 hours at Site A to approximately 68 hours at Site E. Building plumbing samples included first draw, five-minute flush, and full flush cold-water samples. Samples were analyzed for routine physicochemical parameters, heterotrophic plate counts (HPC), and the presence of NTM, which were characterized using plate culture with MALDI-TOF MS for colony identification, qPCR, and genome resolved metagenomics. The onset of the COVID-19 pandemic and associated building closures occurred during the study, impacting water use at three of the buildings (Sites A, B, and E), facilitating an investigation of the impact

of extended low water use on NTM and building water quality. The influence of stagnation was investigated through the collection of samples from buildings with various levels of flushing.

Results

NTM plate culture and MALDI-TOF MS

Presumptive NTM plate culture results ranged from less than the limit of detection (LOD) of 1 colony forming unit (CFU) per volume to 7.3×10^4 CFU/L (Figure S1). LOD values ranged from 1 CFU/L to 333 CFU/L based on the sample volume. Median results in the source waters and through the WTP ranged from 1 CFU/L in the ozone effluent (n=7) to 48 CFU/L in the filter effluent (n=8). In the finished water, the median was 4 CFU/L (n=12). The highest results occurred in the first draw samples at Sites A (median: 1.2×10^4 CFU/L, n=7) and B (median: 1.2×10^4 CFU/L, n=7). The median first draw results in the other buildings ranged from 51 CFU/L at Site D (n=6) to 62 CFU/L at Site C (n=7). Five-minute flush and full flush sample median results ranged from 8 CFU/L in the Site E five-minute flush samples (n=8) to 2.4×10^2 CFU/L in the Site B five-minute flush (n=7).

A subset of the colonies from the NTM culture plates (n=322), representing 47 of the samples, were identified using MALDI-TOF MS. Of these isolates, 60% (n=194) were identified as *Mycobacterium*, 34% (n=108) were not identified, indicating that they were other bacterial or fungal species not represented in the available spectral databases, and 6% (n=20) were identified as other genera including *Bacillus*, *Paenibacillus*, *Brevibacillus*, and *Micromonospora*, which all consist of endospore-forming bacterial species (Table 1) (74–77). Within the isolates identified as NTM, ten species or groups were identified, including *Mycobacterium arupense*, *Mycobacterium asiaticum*, *Mycobacterium aurum*, *M. avium*, *Mycobacterium chelonae* complex, *Mycobacterium franklinii*, *Mycobacterium gordonae*, *Mycobacterium llatzerense*, *Mycobacterium mucogenicum/phocaicum* group, and *Mycobacterium peregrinum*. Of the recovered NTM species, *M. avium* and *M. chelonae* complex are of particular clinical relevance based on prevalence of infections (55, 78, 79).

Table 1. Results of matrix-assisted laser desorption ionization – time of flight mass spectrometry (MALDI-TOF MS) analysis of the plate culture isolates.

Category	Identification	Number of Isolates (% of Total)	Score Range (Median)
Mycobacterium	*Mycobacterium arupense*	2 (<1%)	1.92 – 1.96
	Mycobacterium asiaticum	1 (<1%)	1.61
	Mycobacterium aurum	4 (1%)	1.98 – 2.27 (2.13)
	Mycobacterium avium	8 (2%)	1.99 – 2.20 (2.09)
	Mycobacterium chelonae complex	84 (26%)	1.62 – 2.12 (1.88)
	Mycobacterium franklinii	19 (6%)	1.60 – 1.98 (1.77)
	Mycobacterium gordonae	2 (<1%)	1.81 – 2.01
	Mycobacterium llatzerense	1 (<1%)	1.84
	Mycobacterium mucogenicum/*phocaicum* group	65 (20%)	1.82 – 2.39 (2.11)
	Mycobacterium peregrinum	8 (2%)	2.10 – 2.31 (2.27)
Non-NTM	*Bacillus* spp.	11 (3%)	1.70 – 2.34 (1.94)
	Brevibacillus spp.	6 (2%)	1.77 – 2.40 (1.96)
	Micromonospora sp.	1 (<1%)	1.73
	Paenibacillus spp.	2 (<1%)	1.82 – 2.00
Not identified	--	108 (34%)	0.93 – 1.68 (1.21)

MALDI-TOF MS results showed that none of the isolates recovered from the source waters and WTP ozone effluent and selected for identification were *Mycobacterium* spp. (Figure 1). In the river samples, River-Source isolates included three *Bacillus* spp. isolates and one isolate that could not be identified. For the River-Plant isolates, a larger fraction of isolates could not be identified (55%, n=10), and the remaining isolates were identified as *Bacillus* spp. (28%, n=5), *Paenibacillus* spp. (11%, n=2), and *Micromonospora* sp. (6%, n=1). All isolates from the well samples (Well-Source and Well-Plant) were microorganisms that could not be identified (n=46). Of the two isolates from the ozone effluent samples, one could not be identified, and the other was identified as a *Bacillus* sp.

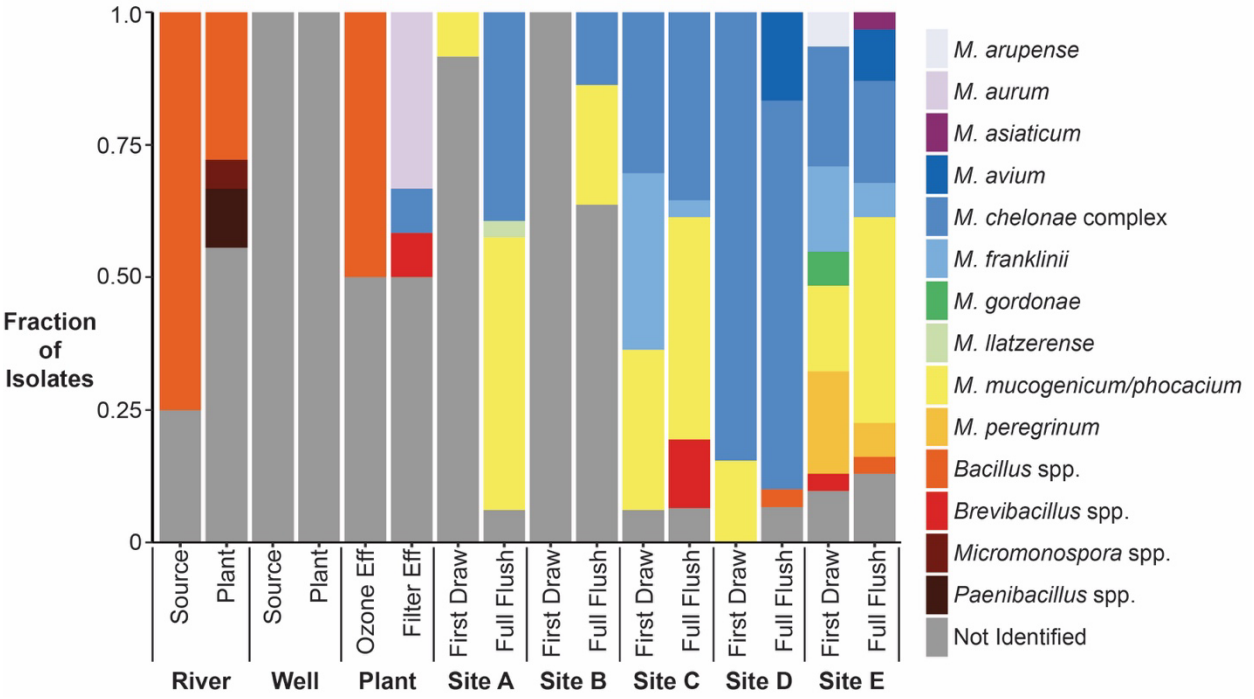

Figure 1. Results of matrix-assisted laser desorption ionization – time of flight mass spectrometry (MALDI-TOF MS) analysis for isolates from the NTM culture plates. No finished water isolates were recovered for the months analyzed. Not Identified: Sample spectra did not match spectral databases.

NTM isolates were identified in the WTP filter effluent and all downstream sampling locations, except the first draw samples from Site B. In the filter effluent, isolates included *M. aurum* (n=4), *M. chelonae* complex (n=1), *Brevibacillus* sp. (n=1), and six isolates that could not be identified. No isolates from the finished water were analyzed using MALDI-TOF MS due to the low number of isolates recovered. In the distribution system, the majority of isolates (68%, n=28 of 41) were NTM, with *M. chelonae* complex (n=8) and *M. mucogenicum/phocaicum* group (n=8) being the most frequently recovered. At Sites A and B, the fractions of NTM isolates recovered from first draw samples were low (<10%) but increased in the full flush samples (>35%). Increased recovery and diversity of NTM were observed at higher water age sites. At Sites D and E, the fractions of isolates identified as NTM were 93% (n=40 of 43) and 85% (n=53 of 62), respectively. The highest diversity in NTM isolates was observed at Site E, where eight different NTM species or groups were identified (*M. arupense*, *M. asiaticum, M. avium*, *M. chelonae* complex, *M. franklinii*, *M.*

gordonae, *M. mucogenicum/phocaicum* group, and *M. peregrinum*). *M. avium* was only isolated from the full flush samples from Site D (n=5) and Site E (n=3).

The fraction of isolates identified as NTM for each sample analyzed using MALDI-TOF MS was used to adjust the presumptive NTM plate culture results, yielding the adjusted plate counts (Figure S2). Although the highest presumptive CFU/L results occurred in the first draw samples from Sites A and B, adjusted NTM CFU/L results were lower due to the large fraction of non-NTM isolates. The sample with the highest adjusted NTM plate count was the July Site C first draw sample at $5.6x10^2$ CFU/L. As most isolates picked from the Site C, D, and E culture plates were identified as NTM, the adjusted plate counts were generally similar to the presumptive NTM counts.

NTM quantification using qPCR

Similar to the plate culture results, the highest concentrations of NTM using qPCR were observed in the first draw samples from Sites A and B (Figure 2). The lowest NTM gene copy concentrations were found in the Well-Source and the Ozone Effluent, where most samples were below the limit of quantification (LOQ) of 41 gene copies per reaction (gc/rxn). LOQs in gene copies per liter (gc/L) ranged from $2x10^2$ to $3x10^5$ gc/L and varied based on the sample volume and qPCR dilution factor. NTM gene copy concentrations in the river samples were higher, with a median of $7.3x10^4$ gc/L (n=11). Gene copy concentrations increased after filtration, from less than the LOQ in the Ozone Effluent to a median of $1.1x10^4$ gc/L in the Filter Effluent (n=7). In the Finished Water samples, the median NTM gene copy concentration was $1.4x10^3$ gc/L (n=8). At Sites A and B, maximum NTM gene copy concentrations in the first draw samples reached $5.5x10^5$ gc/L and $4.3x10^7$ gc/L, respectively.

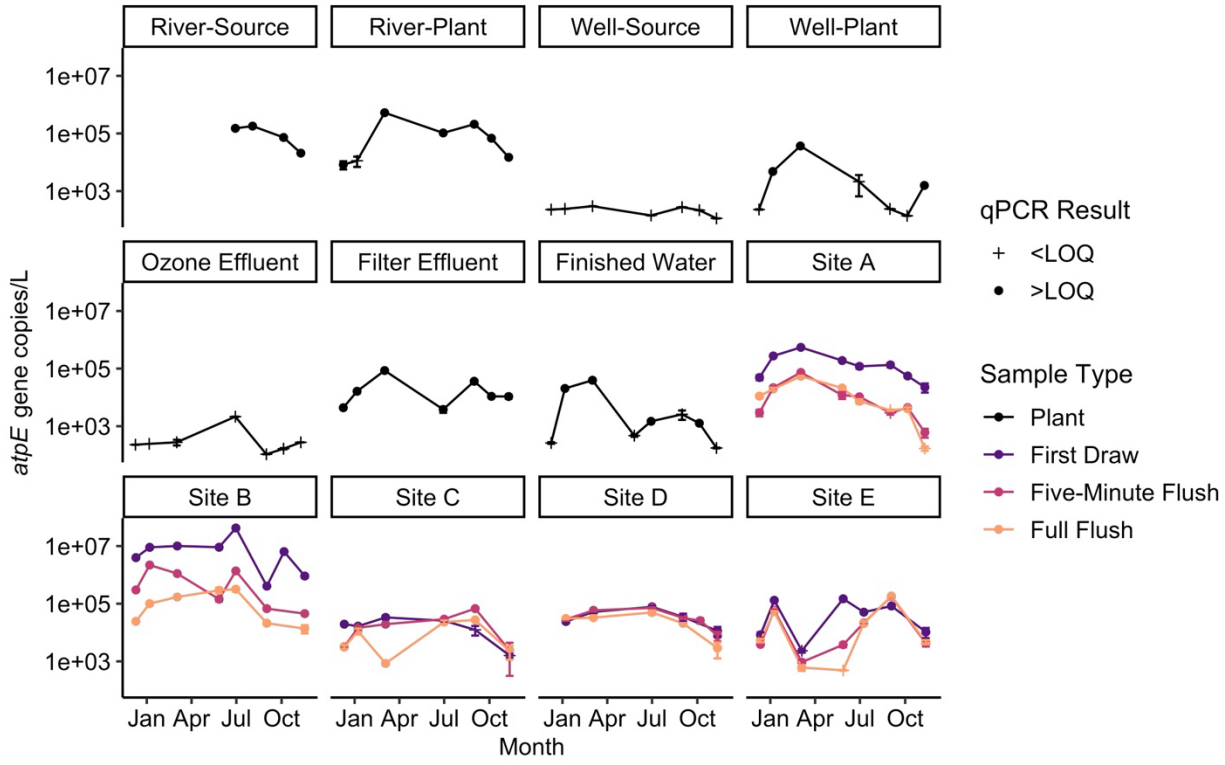

Figure 2. NTM *atpE* gene copy concentrations in all samples analyzed over the study period. Values below the LOQ (crosses) were set to one-half the LOQ and then converted to gene copies per liter. Values above the LOQ are shown as circles. Error bars show one standard deviation above and below the mean of triplicate qPCR reactions. Source: Based on data from Dowdell et al., 2022 (43).

In the building plumbing samples, the highest NTM gene copy concentrations were observed in the first draw samples (median: 6.7×10^4 gc/L, n=34), followed by the five-minute flush samples (median: 2.4×10^4 gc/L, n=34) and full flush samples (median: 2.1×10^4 gc/L, n=33). NTM gene copy concentrations in the full flush samples, which captured distribution system water quality at the building locations, ranged from a median of 5.3×10^3 gc/L at Site E to 1.0×10^5 gc/L at Site B. Differences in NTM gene copy concentrations were observed over the sampling months. The months with the lowest median NTM gene copy concentrations in first draw samples were November (1.1×10^4 gc/L, n=5) and December (3.4×10^4 gc/L, n=4), while the highest median NTM gene copy concentrations in the first draw samples were found in October (3.3×10^6 gc/L, n=2), May (1.9×10^5 gc/L, n=3), and January (1.3×10^5 gc/L, n=5). In the full flush samples, the months with the highest median NTM gene copy concentrations were March (3.3×10^4 gc/L, n=5) and

January ($3.0x10^4$ gc/L, n=5), and the month with the lowest median NTM gene copy concentration was November ($2.9x10^3$ gc/L, n=5).

NTM species identified by metagenomic analysis

Processing of the metagenomic sequences from the WTP filter effluent, finished water, and full flush samples yielded five high quality (>50% completeness and <10% redundancy) NTM metagenome-assembled genomes (MAGs) (Figure 3, Table S4). Of these, four were identified to the species level, representing *M. phocaicum*, *M. llatzerense*, *M. gordonae,* and *M. arupense*. The MAG with the highest relative abundance across all samples was the one identified as *M. gordonae*, which occurred at a relative abundance of 0.91 reads per kilobase million (RPKM) in the December Site B full flush sample. The *M. gordonae* MAG was also detected in the Site B full flush samples from the other months, with relative abundances ranging from 0.08 to 0.51 RPKM, and in the October Site E full flush sample (0.15 RPKM). The NTM MAG was detected in the full flush samples from Site A in December, Site B in March, July, and October, and Site E in October, with relative abundances ranging from 0.04 to 0.14 RPKM. The only detection of the *M. phocaicum* MAG was in the December Site B full flush sample (0.49 RPKM). The highest relative abundance of the *M. arupense* MAG was in the July Site B full flush sample (0.65 RPKM). The *M. arupense* MAG was detected in the other Site B samples, as well as the Site A samples from December and October and the Site E samples from July and October. The *M. llatzerense* MAG was detected in the December Site B full flush sample (0.36 RPKM) and October Site E full flush sample (0.17 RPKM). No reads corresponding to the filter effluent or finished water samples met the 25% coverage minimum.

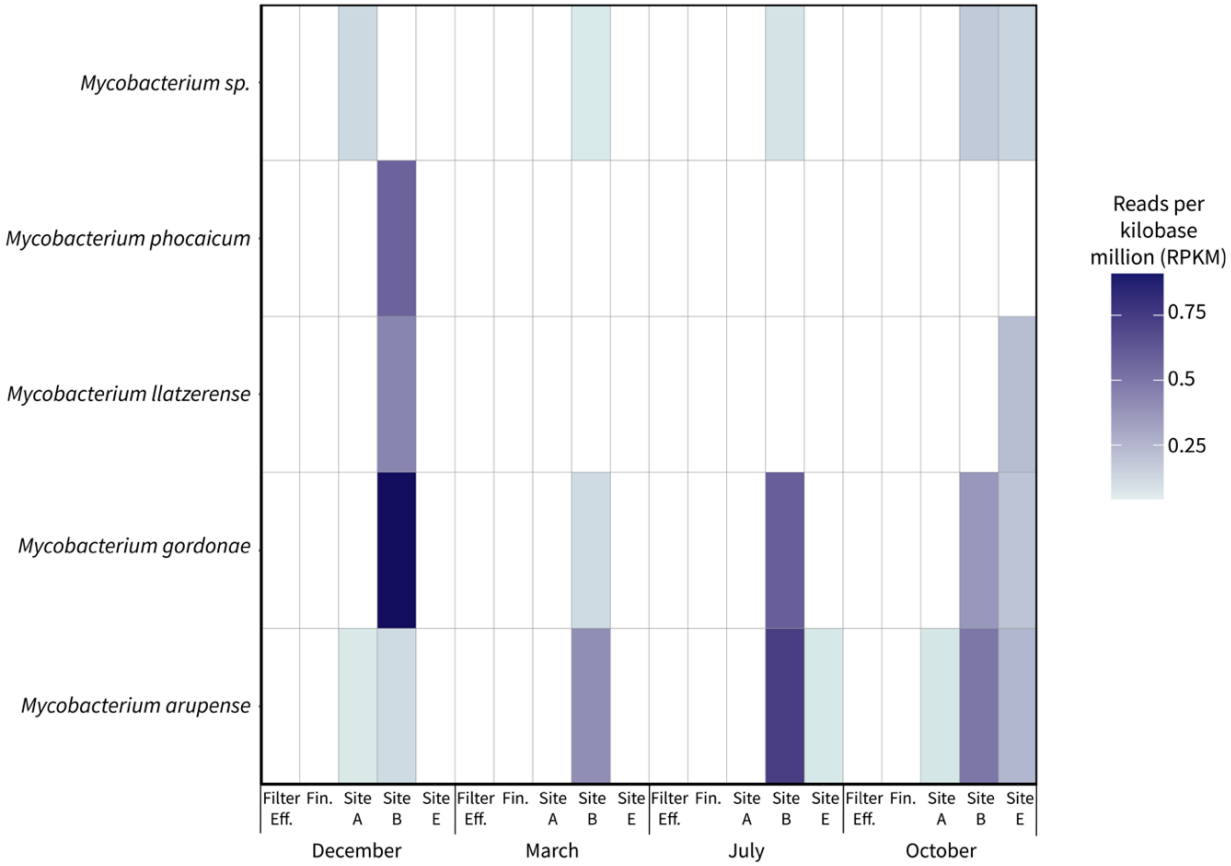

Figure 3. Metagenome-assembled genomes (MAGs) identified as *Mycobacterium*. Relative abundances of reads mapping to the NTM MAGs are represented by reads per kilobase million (RPKM). Filter Eff.: Filter Effluent; Fin.: Finished Water, Site A: Site A full flush sample, Site B: Site B full flush sample, Site E: Site E full flush sample.

The highest relative abundances for all the MAGs except for the *Mycobacterium* sp. MAG occurred in the Site B full flush samples. However, the months with the highest observed relative abundances varied, with the highest relative abundances for the *M. phocaicum*, *M. llatzerense*, and *M. gordonae* MAGs occurring in December, while the highest relative abundance of the *M. arupense* MAG occurred in July. The December Site B full flush sample was also one of only two samples in which four of the five MAGs were detected. The other sample in which four of the MAGs were detected was the October Site E full flush sample. While most of the MAGs were detected across seasons, *M. phocaicum* was only detected in December and *M. llatzerense* was only detected in December and October.

Physicochemical and HPC analyses

Water sample temperatures fluctuated with the season, with temperatures ranging from less than 5°C in February 2020 to approximately 23°C in August 2020 (Figure S3). Although the Well-Source temperatures were relatively stable over the year (median: 11.9 ± 0.6°C, n=9), Well-Plant temperature trends were influenced by season (median: 18.0 ± 5.0°C, n=9). pH values did not vary substantially at any of the sampling locations over the sampling campaign (standard deviations ranged from <0.1 - 0.2; Figure S4). Median pH values were 7.3 in the well samples (Well-Source and Well-Plant; n=18), 8.1 in the river samples (River-Source and River-Plant; n=12), and 9.3 across all other sampling locations (n=175). Turbidities in the pre-filtration samples were typically greater than one nephelometric turbidity unit (NTU), with the highest turbidities occurring in the well water samples (median: 18.0 NTU, n=18) and the WTP ozone effluent (median:7.0 NTU, n=9; Figure S5). Median turbidity values for samples post-filtration were less than 0.3 NTU at all locations, except the first draw samples from Site E, where the median turbidity was 0.5 NTU (n=10). Monochloramine concentrations in the finished water were approximately 3.0 mg/L across all sampling events (median: 3.0 mg/L as Cl_2, n=11; Figure 4). Building water sample monochloramine concentrations varied by location, sample type, and month. Of the first draw building water samples (median: 2.42 mg/L as Cl_2, n=47), the lowest monochloramine concentrations were observed in the May and June samples at Site B, which were near or below the LOD of 0.04 mg/L as Cl_2.

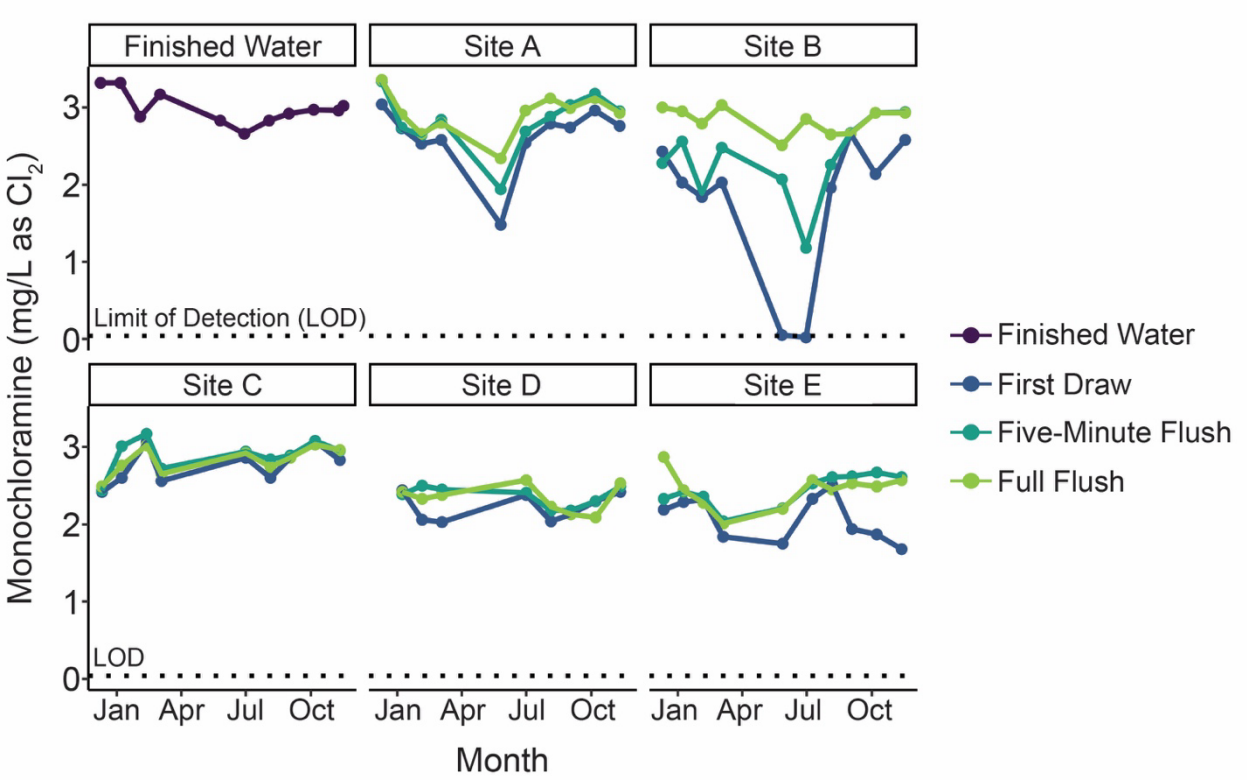

Figure 4. Monochloramine concentrations in the finished water and the five distribution system sites (Sites A – E). Samples at the distribution system sites included first draw, five-minute flush, and full flush. The dotted line indicates the limit of detection (LOD, 0.04 mg/L as Cl_2). Source: Based on data from Dowdell et al., 2022 (43).

HPC results were generally low after secondary disinfection, except in some first draw samples (Figure S6). River water HPC results ranged from $1.5x10^3$ to $6.3x10^3$ CFU/mL (median: $3.9x10^3$ CFU/mL, n=7). HPC results were lower in the well water samples, which ranged from <1 CFU/mL to $8.3x10^2$ CFU/mL (median: $6.4x10^2$ CFU/mL, n=16). Notably, the HPC results were significantly higher in the Well–Plant samples compared to the Well–Source samples ($p<0.05$), indicating bacterial growth during transmission of the well water to the WTP. Ozone effluent HPC results ranged from 5 to $1x10^3$ CFU/mL (median: $1.2x10^2$ CFU/mL, n=9). Filter effluent HPC results were generally higher than in the ozone effluent, ranging from 18 to $2.6x10^4$ CFU/mL (median: $1.0x10^4$ CFU/mL, n=9). HPC results in the finished water ranged from 1 to 30 CFU/mL (median: 4 CFU/mL, n=11). In the distribution system, the full flush samples had the lowest HPC results overall (median: 8 CFU/mL, n=45) while the first draw samples were the highest (median: $1.7x10^2$ CFU/mL, n=45). However, median HPC results for first draw samples varied by site, with

median results at Sites A ($4.0x10^3$ CFU/mL, n=9), B ($4.4x10^3$ CFU/mL, n=10), and E (median: $1.8x10^2$ CFU/mL, n=10) being at least one order of magnitude higher than the median HPC results at Sites C (41 CFU/mL, n=8) and D (4 CFU/mL, n=8).

Impacts of decreased water use and flushing on physiochemical parameters and NTM

The COVID-19 pandemic occurred during sampling, resulting in lower than usual water usage in several distribution system monitoring locations beginning in mid-March 2020 (Figure 5). At Site A, the monthly water use in March, April, and May 2020 was at least 46% lower than in 2019, returning to values similar to 2019 beginning in June 2020. At Site B, March 2020 water use decreased to 48% of 2019 use, with the lowest water use in April and May 2020, both of which were at least 90% less than the monthly water use in 2019. At Site B, water use rebounded beginning in July 2020. At Site E, monthly water use in 2020 was below 2019 levels from March through May 2020, but then was substantially higher in the remaining months of 2020.

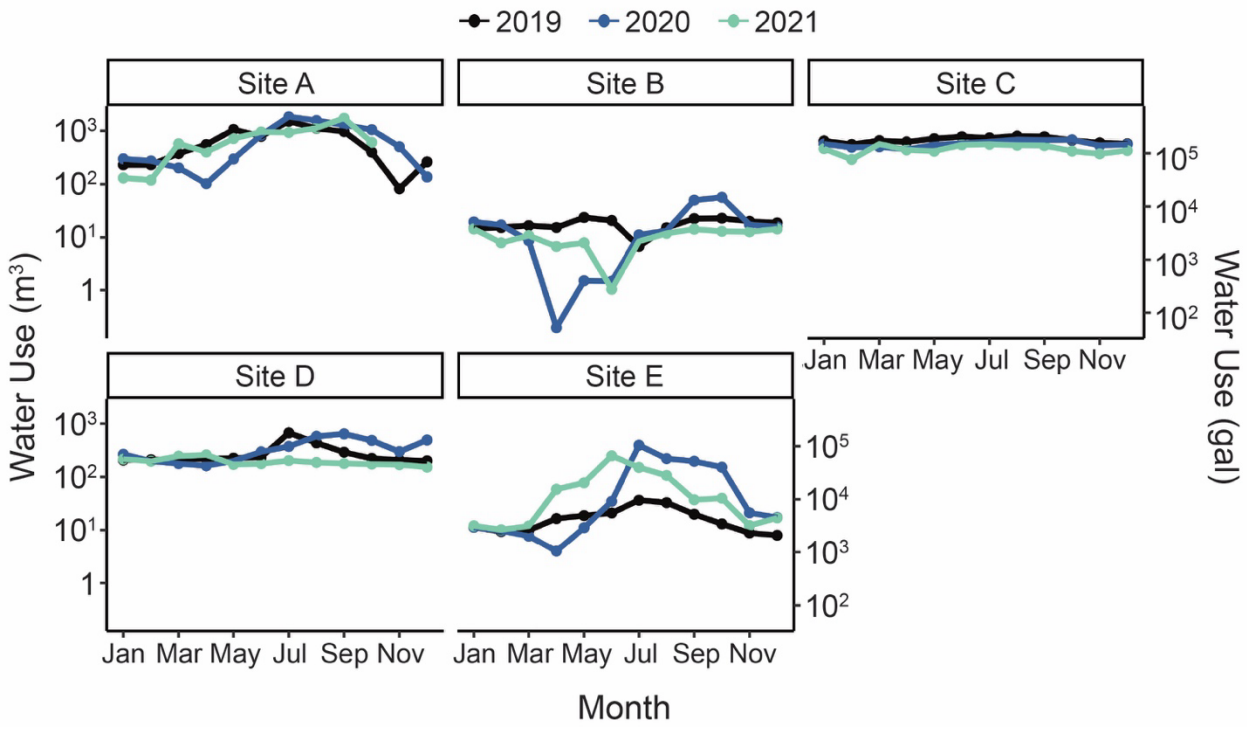

Figure 5. Monthly building water usage in the years 2019, 2020, and 2021. Units are cubic meters (left y-axis) and gallons (right y-axis).

The low water use at Sites A, B, and E, was associated with decreased disinfectant residual (Figure 4). Specifically, the monochloramine residuals in the first draw samples from Sites A and B in May 2020 were at least 40% lower than the monochloramine residuals in the March samples, which were collected prior to building closures. At Site B, first draw monochloramine residuals were near or below the detection limit (0.04 mg/L) in both May and July. At Sites A and B, first draw samples during the low water use period also had higher turbidities than the first draw samples once water use increased to normal levels (Figure S5).

Tap flushing improved water quality parameters at several building plumbing locations, though the degree of improvement and amount of flush time required for improved conditions varied by building and parameter. Flushing significantly reduced NTM qPCR results at Sites A and B (first draw versus full flush, $p<0.05$) but did not significantly reduce NTM qPCR results at Site C, D, and E. The flush duration required to reduce NTM qPCR results at Sites A and B varied. At Site A, the NTM concentrations in the five-minute flush and full flush were similar, suggesting that distribution system water was reached after five minutes flushing, or that building plumbing NTM levels were similar to distribution system levels. In contrast, flushing beyond five minutes reduced NTM gene copy concentrations by an average of 0.7 log at Site B.

Influence of water age and physicochemical parameters on NTM

To investigate the influence of distribution on NTM and physicochemical parameters, correlation analysis was performed. The correlation analysis included temperature, pH, turbidity, monochloramine concentration, estimated water age at the site, and the qPCR results for the full flush samples. As this correlation analysis was focused on distribution system water age, first draw and five-minute samples were not included. The results showed that estimated water age had a strong negative correlation with monochloramine concentration (Kendall's tau: -0.50, $p<0.01$; Figure S7). Estimated water age was also positively correlated with turbidity, though the association was weaker (Kendall's tau: 0.24, $p>0.05$). The total NTM concentration was not strongly associated with any of the water quality parameters, though a weak positive correlation was observed with pH (Kendall's tau: 0.23, $p>0.05$).

Discussion

375 This study is unique in that monitoring had started several months prior to the COVID-19
376 pandemic, which allowed for evaluation of how decreased water use in buildings impacted general
377 water quality parameters and NTM. At Sites A and B, where buildings were closed beginning in
378 mid-March 2020, lower water qualities were observed in first draw samples, including turbidities
379 as high as 4.6 NTU, HPC results as high as $8.0x10^4$ CFU/mL, and one sample without a measurable
380 monochloramine residual. In addition to general water quality impacts, the concentrations of NTM
381 gene copies were also higher, with values as high as $4.3x10^7$ gc/L. These NTM concentrations
382 were similar to those reported in another study that investigated the impact of pandemic-related
383 building closures on NTM using qPCR (16). As the buildings re-opened and water use increased
384 in late summer and fall 2020, the NTM concentrations and physicochemical parameters, including
385 monochloramine residual, generally improved. In Sites C and D, which remained fully open during
386 the pandemic lockdown, water use was maintained as usual. This continuation of similar water use
387 patterns was reflected in the consistency of the NTM concentration and monochloramine residuals
388 across the year of sampling. Although most studies investigating the impacts of COVID-19
389 pandemic-related building closures focused on other parameters, such as metals or *L. pneumophila*
390 (32–34, 80), this study shows the importance of also considering the impact of low water use on
391 NTM occurrence.
392
393 By sampling from drinking water sources to taps in buildings in the distribution system, this study
394 investigated how NTM populations change through water treatment and delivery. In well water,
395 NTM gene copy concentrations changed even prior to treatment, increasing during conveyance
396 from the well to the WTP. Transmission of well water to the WTP also significantly increased
397 water temperatures ($p<0.05$; Figure S3) and HPC results ($p<0.05$, Figure S6) in the Well-Plant
398 samples compared to the Well-Source samples. As a result, the temperature curve for the Well-
399 Plant samples is more similar to that of the surface water samples than the Well-Source samples
400 (Figure S3). While ozonation effectively reduced NTM gene copy concentrations to below the
401 LOQ, these concentrations rebounded after filtration (median:10^4 gc/L) and only decreased after
402 secondary disinfection in certain months. These findings support previous studies, which observed
403 that NTM concentrations increased after filtration (25, 81). Additionally, these results are
404 consistent with the importance of filtration in shaping bacterial communities in the finished water
405 (82).

406

407 Plate culture analyses indicated that the WTP samples were dominated by non-NTM, including
408 spore-forming bacteria and microorganisms that could not be identified. The first draw samples
409 from Sites A and B and the full flush sample from Site B also yielded either all or a majority of
410 non-NTM isolates. The WTP filter effluent was the first location where NTM isolates were
411 recovered from culture plates and, except for Site A and B first draw samples, the building
412 plumbing sample isolates were at least 75% NTM. The low recovery of NTM isolates prior to
413 filtration is likely due to high concentrations of non-NTM microorganisms in those samples. In
414 previous studies focused on NTM quantification in source waters or varied water types, antibiotics
415 were generally added to the NTM culture media in addition to using decontamination strategies
416 (23, 38). Therefore, only using decontamination, as was done in this study, was insufficient for the
417 recovery of NTM in certain samples with high concentrations of non-target microorganisms.
418 However, the selected culture method and similar methods utilizing cetylpyridinium chloride
419 (CPC), a common biocide, as a pre-treatment have been commonly used to quantify NTM in
420 drinking water samples, including building water samples (22, 30, 35), and the method selected in
421 this study performed well at all sites post-disinfection except for the first draw samples from Sites
422 A and B.

423

424 The occurrence of specific species of NTM varied by site, with *M. aurum* only isolated from the
425 WTP filter effluent, *M. llatzerense* only recovered from Site A, *M. avium* only recovered from
426 Sites D and E, and *M. arupense, M. asiaticum*, and *M. gordonae* only isolated from Site E. The
427 finding that *M. avium* was only recovered from the sites with the highest water ages supports the
428 results of a previous study, which reported that the relative abundance of *M. avium* increased with
429 water age (14). Other NTM species, including *M. chelonae* complex, *M. mucogenicum*/*phocaicum*
430 group, and *M. franklinii* were isolated from across sites with varying water ages. Overall, the NTM
431 isolates were predominantly *M. chelonae* complex (43%, n=84) and *M. mucogenicum/phocaicum*
432 group (34%, n=65). A recent study investigating NTM species in chloraminated distribution
433 systems also reported that *M. mucogenicum/phocaicum* group was the most common species
434 recovered, though in that study the proportion was higher at 76% of the isolates (21).

435

436 Among the NTM isolates recovered in this study, *M. avium* and *M. chelonae* complex are most
437 associated with human infection (83). While *M. avium* was only isolated from the full flush
438 samples from Site D (n=5) and Site E (n=3), *M. chelonae* complex was the most commonly
439 recovered isolate, representing 26% (n=84) of all isolates and 43% of NTM isolates. Though not
440 as commonly associated with disease as *M. avium* and *M. chelonae*, *M. mucogenicum* and *M.*
441 *peregrinum* have also been reported to cause human infections (57, 84, 85). The prevalence of
442 clinically-relevant NTM species other than *M. avium* emphasizes the importance of monitoring for
443 additional NTM species, such as *M. abscessus* and *M. chelonae*, when investigating NTM
444 occurrence in drinking water. Studies have also reported geographical differences in the NTM that
445 account for the majority of infections (49, 56), which should be considered when selecting species-
446 specific methods for NTM quantification.
447
448 Metagenomic analysis detected four species of NTM in the samples analyzed: *M. phocaicum*, *M.*
449 *gordonae*, *M. llatzerense*, and *M. arupense*, all of which were also recovered by plate culture.
450 However, there were differences in where these NTM species were detected when using plate
451 culture versus metagenomic analysis. *M. phocaicum*, which was detected in the December Site B
452 full flush sample by metagenomic analysis, was isolated by plate culture from the filter effluent
453 and all distribution system sampling sites. *M. arupense*, which was detected at Sites A, B, and E
454 by metagenomic analysis, was only recovered from Site E with plate culture. Similarly, *M.*
455 *gordonae* was detected in Site B and E samples by metagenomic analysis but only in Site E by
456 plate culture, and *M. llatzerense* was detected in Sites B and E using metagenomic analysis but
457 only in Site A by plate culture. Therefore, both methods detected certain species that were not
458 identified using the other method at particular sites. Possible reasons for this disagreement include
459 that the sequencing depth may have been insufficient to detect rarer species, that not all microbial
460 community diversity was captured during binning, that the smaller sample volumes associated
461 with the plate culture method may have prevented detection of rarer species, and that the plate
462 culture method may have de-selected for certain species. These results highlight the benefits of
463 using both cultivation-dependent and -independent methods, as it allows for a more complete
464 understanding of NTM occurrence.
465

Although the culture-based method resulted in the identification of a larger number of NTM species, there are known biases associated with NTM plate culture methods. The pre-treatment step, for example, has been shown to impact species and strains of NTM differently, with one study reporting that a dose of 0.1% CPC for 30 minutes reduced culturable water-grown *M. avium A5* by approximately 80% but only reduced *M. avium Va14 (T)* by 20% (36). Differences in susceptibility of NTM species to pre-treatments by growth category (rapid-growing or slow-growing) have also been observed, with one study reporting that rapidly growing NTM species such as *M. fortuitum*, *M. abscessus*, and *M. peregrinum* may be more sensitive to decontamination (37). However, it is difficult to balance the need to reduce colony formation by non-NTM with recovery of target species. A new NTM plate culture medium that can be used without sample pre-treatment has been described recently (86). This method may overcome some of the limitations of existing methods, though additional comparisons are needed to validate its selectivity (86, 87).

Sampling over the course of a year allowed for investigation of the impacts of temperature and season on NTM populations. While strong seasonal trends were not observed, the NTM gene copy concentrations in full flush samples across all sites were lowest in November (median: 2.9×10^3 gc/L, n=5) and highest in March (3.3×10^4 gc/L, n=5). The impact of season could not be addressed for the first flush and five-minute flush samples due to the changes in water use patterns at several of the sites due to the COVID-19 pandemic. The impact of season on NTM species occurrence also could not be determined based on the MADLI-TOF MS results due to the low numbers of isolates analyzed per site per month and the strong influence of site on NTM species. The metagenomic analysis found the highest relative abundances of NTM MAGs in December and July. Overall, low water use appeared to be the driving factor of NTM concentration in building plumbing rather than season. While few studies have investigated the impact of season on NTM concentrations, one previous study reported higher concentrations of NTM gene copies in a distribution system during the summer compared to the winter (88). However, another study reported significantly higher concentrations of NTM in distribution systems in winter and spring, which it attributed to lower water use (19). Due to COVID-19 pandemic-associated building closures during this study, it is difficult to determine whether the generally higher NTM concentrations observed in the summer are linked to season or were the result of shifts in water

use patterns. Additional studies are needed to further investigate the potential impacts of both water use and season on NTM in drinking water distribution systems.

Methods

Source water and treatment plant

The full-scale WTP sampled in this study has been described previously (15, 27, 82). Briefly, the City of Ann Arbor WTP is a 50 million gallon per day ($1.9x10^5$ m^3/day) facility in Ann Arbor Michigan, USA that treats groundwater and river water, which are blended at a ratio of approximately 1:6, using two-stage excess lime softening, coagulation, flocculation, sedimentation, ozonation, biological filtration, and monochloramine disinfection (3 mg/L as Cl_2). Ozonation is used to achieve 0.5-log removal of *Cryptosporidium*. Typical ozone concentration × time (CT) values range from 0.3 to 1 mg/L. The biological filters contain either dual media sand-granular activated carbon (GAC) or GAC alone. In June 2020, the WTP began using an ultraviolet (UV) disinfection system intermittently, which was intended to provide additional *Cryptosporidium* inactivation when the plant is operated with single-stage lime softening during maintenance periods. The UV system was in operation during the June, August, and September sampling events.

Sample collection

Samples were collected approximately monthly from December 2019 through November 2020. Samples were not collected due to the COVID-19 pandemic in April 2020, and only a partial sampling event (finished water reservoir [Finished Water] and three buildings in the distribution system [Sites A, B, and E]) occurred in May 2020 due to building access restrictions. Samples were collected using sterile polypropylene bottles. All source water and treatment plant samples except the Ozone Effluent samples were either collected from the sampling taps that were continuously flowing (River-Plant, Well-Plant, and Finished Water) or were collected after flushing the taps for at least five minutes (River-Source, Well-Source, and Filter Effluent). Source water samples were collected from sampling taps at the source water pump stations (River–Source and Well–Source) as well as from the sampling taps at the WTP used for compliance monitoring (River–Plant and Well–Plant). River sample volumes ranged from 2 to 4 L and well samples ranged from 10 to 15 L. The Ozone Effluent samples were collected from a clearwell after the

ozone contactors and prior to filtration. Ozone Effluent sample volumes ranged from 10 to 20 L. Filter Effluent samples were the combined effluents of six filters and ranged in volume from 10 to 20 L. Finished Water samples were collected from a sampling tap in the laboratory of the WTP used for compliance monitoring.

Cold water samples were collected from sinks in five buildings, Sites A – E, receiving water from the WTP via the distribution system (Table S1). These locations are a subset of the 13 used for distribution system compliance monitoring. Sites were selected to capture the full range of estimated water ages at compliance sites, as determined using an EPANET distribution system model (89). Sites were ordered based on estimated water age and selected to capture a range of water quality characteristics based on historical data provided by the WTP, which included concentrations of total chlorine and nitrogen species, and HPC. Additional details on the distribution sampling locations are included in Table S1. Three samples were collected from each distribution sampling location: 1) a 2 L first draw sample, representing the water in the fixture and building plumbing line supplying the fixture, 2) a 10 L five-minute flush sample, meant to represent the typical monitoring sample collected by the utility, and 3) a 10 to 20 L fully flushed sample, meant to represent distribution system water at the location. To mimic the procedure for compliance sampling, the tap was not disinfected for the first draw samples but was sprayed with a 10% bleach solution prior to collection of the five-minute flush and full flush samples. For the full flush samples, temperature and pH were monitored and samples were collected when successive readings stabilized.

Physicochemical analyses

Temperature, pH, conductivity, and total dissolved solids (TDS) were analyzed immediately on-site using a combination probe (Hanna Instruments HI98121, Smithfield, RI, USA). Total chlorine, free chlorine, and monochloramine were also analyzed on-site using a portable colorimeter (Hach DR900, Loveland, CO, USA) and powder pillows (Hach Methods 10250, 10245, and 10200). Samples for total organic carbon (TOC) were transferred to carbon-free glass vials and acidified with 85% phosphoric acid. TOC samples were analyzed with the non-purgeable organic carbon (NPOC) method using a Shimadzu TOC analyzer (TOC-V, Japan). Turbidity was measured using a benchtop turbidimeter (Hach TU5200).

558

559 ***DNA collection and extraction***

560 Water samples were filtered onto sterile 0.22 µm copolymer cartridge filters (EDM Millipore,
561 USA, cat. no. SVGPL10RC) using sterile tubing (Masterflex, USA) and peristaltic pumps. Filters
562 were then placed in sterile bags. Samples collected from December 2019 through March 2020
563 were filtered in the field and flash frozen using dry ice and ethanol, then transported to the
564 laboratory and frozen at -80°C. Due to COVID-19 pandemic restrictions, samples collected after
565 March 2020 were transported in coolers with cold packs to the laboratory for filtration and were
566 placed in the -80°C freezer rather than flash freezing with dry ice.

567

568 DNA was extracted using a modified form of the QIAGEN DNeasy PowerWater kit (QIAGEN,
569 Hilden, DEU, cat. no. 14900-100-NF) method, which includes additional enzymatic (Proteinase K
570 and lysozyme) and chemical (chloroform-isoamyl 24:1) lysis steps (90). DNA purity was
571 measured using a Nanodrop 1.0 spectrophotometer (Thermo Fisher Scientific, Waltham, MA,
572 USA) and DNA concentrations were measured using the Qubit™ dsDNA High-Sensitivity assay
573 kit (Thermo Fisher Scientific, cat. no. Q32851).

574

575 ***NTM and HPC culture methods***

576 Samples for HPC analyses were transferred to 100 mL sterile vessels containing excess sodium
577 thiosulfate and analyzed using the pour-plate method (SM 9215-B-2000) (91). NTM plate culture
578 was performed for all sampling events except February, May, and August 2020. Sample volumes
579 for NTM plate culture ranged from 3 mL to 1 L and were adjusted each month based on colony
580 counts for the previous month in an effort to maximize the number of plates that yielded 1 to 300
581 colonies (Table S2). Samples (n=148) were processed in duplicate except when insufficient
582 volume was available to process in duplicate (occurred in 13 cases, all of which were first draw
583 samples with limited volume). Filter concentration followed by pre-treatment with CPC was
584 employed. Briefly, samples were transferred to sterile glass bottles and dosed with a sterile 2%
585 CPC solution to a final concentration of 0.04%. Samples were incubated at room temperature for
586 30 minutes, then immediately filtered onto 0.45 µm filters (Thermo Fisher Scientific, cat. no. 09-
587 719-555) using sterile glass filtration bowls and pedestals and a vacuum filtration manifold (EDM
588 Millipore, USA) in a fume hood. After filtration, the filters were rinsed with 50 mL of sterile

ultrapure water (Invitrogen™ UltraPure™ DNase/RNase-Free Distilled Water, Thermo Fisher Scientific, cat. no. 10977015) to remove residual CPC. Filters were plated onto sterile Middlebrook 7H11 agar (Thermo Fisher Scientific, cat. no. R454002) supplemented with oleic albumin dextrose catalase (OADC, BBL™ Middlebrook OADC Enrichment, BD Biosciences, Franklin Lakes, NJ, USA, cat. no. 212351) using sterile tweezers in a biological safety cabinet. Plates were allowed to dry in the biological safety cabinet, then sealed with Parafilm (Bemis Company, USA), inverted, and placed in loosely sealed plastic bags. Plates were incubated in the dark at 35°C for at least one month and observed at least once per week. Controls were included for at least every sampling day and included plate controls (agar plates only, n=24) and filtration controls (50 mL of ultrapure water filtered and plated, n=26).

Plates were classified into four groups, either negative (no colonies), positive and quantifiable (CFU between 1 and 300), too numerous to count (TNTC; CFU greater than 300), or no count (NC). The NC category represented plates that contained spreader colonies or other colonies that grew to overtake the plates, preventing accurate counting of colonies, which may have originated in the samples or may have been the result of contamination. Colony counts from the duplicate plates were averaged to obtain a plate count for each sample. For calculations and plotting, TNTC plate counts were set to 300 CFU and negative plates were set to 0.5 CFU. Excluding controls, a total of 284 plates were processed for 148 samples. Of these, 33 plates (12%) were TNTC, 44 (15%) were negative, and 20 (7%) fell in the NC category, leaving 187 plates (66%) from 107 (72%) samples that were counted. Only four of the 50 controls showed contamination. For all plates that were counted, at least 10% of the colonies were picked for confirmation testing. Colonies were selected to cover a range of morphologies and growth characteristics. Colonies were picked in a biological safety cabinet using sterile pipette tips. Biomass was then transferred to sterile 0.6 mL tubes containing 50 µL of ultrapure water and frozen at -80°C.

NTM isolate protein extraction MALDI-TOF MS analysis

Colonies (n=391) from the NTM culture plates from the December, January, March, June, and October sampling events were analyzed using MALDI-TOF MS. Due to the number of isolates, only isolates from first draw and full flush samples were selected for analysis. Briefly, colonies were thawed and used to inoculate 5 mL tubes containing Middlebrook 7H9 with Tween® 80

(Hardy Diagnostics, Santa Maria, CA, USA, cat. no. C62) and incubated at 35°C. Once growth was observed, isolates were streaked onto Middlebrook 7H11 with OADC and incubated at 35°C. Isolates that failed to grow on Middlebrook 7H11 with OADC at 35°C were streaked onto a Middlebrook 7H11 with OADC and incubated at 32°C and streaked onto an LB agar and incubated at 35°C. Seventy of the isolates (18%), from 25 samples, failed to grow despite the use of multiple temperatures and media and therefore could not be analyzed using MALDI-TOF MS.

For the protein extraction, each isolate was extracted in duplicate by picking one µL loopfuls of colonies from the culture plates and transferring them to two 2 mL tubes containing 200 µL of 0.5 mm zirconium/silica beads and 500 µL of 70% ethanol. Beads were heat-treated prior to use by incubating in a 450°C oven for at least three hours. Tubes were vortexed for 15 minutes, then centrifuged at 10,000xg for two minutes. The supernatant was discarded, and tubes were allowed to dry in a biosafety cabinet. Once dry, tubes were transferred to a fume hood, where 20 µL of 70% formic acid was added to each tube. Tubes were vortexed and incubated for five minutes, then 20 µL of acetonitrile was added. Tubes were centrifuged again at 10,000xg for two minutes and the supernatant was used for MALDI-TOF MS analysis within four hours of extraction. Each protein extraction session included a negative control (an empty bead tube), a non-NTM positive control (colonies of *Escherichia coli*), and an NTM positive control (colonies of *Mycobacterium smegmatis*).

MALDI-TOF MS analysis was performed at the University of Michigan Microbiome Core using a Bruker MALDI Biotyper® sirius system (Bruker Daltonics, Billerica, MA, USA) and a 96 target brushed steel plate. Each MALDI-TOF MS plate analysis began with analyzing two targets of Bacterial Test Standard (BTS, Bruker Daltonics, Billerica, MA, USA, cat. No. 8255343) to ensure the system was functioning properly. All extracts were spotted in duplicate and spectra were analyzed using both the standard instrument library and the *Mycobacterium* genus-specific Mycobacteria RUO Library (Bruker Daltonics). BTS scores were required to be above 2.0 and *M. smegmatis* positive control results were required to be above 1.7. The maximum score between the duplicate extractions and duplicate spots (four results per isolate) was used for identification. Minimum scores for identification followed manufacturer guidance; NTM isolates were required

to score at least 1.6 to be identified to the species or complex level, while non-NTM isolates were required to score at least 1.7 (68). Per the manufacturer, NTM isolates with scores less than 1.6 were not identifiable, isolates with scores between 1.60 and 1.79 were "low confidence" identifications, and isolates with scores of at least 1.8 were considered "high confidence" identifications. Score results by genus and species are provided in Table S3. Due to the inability of the method to distinguish between *M. chelonae*, *Mycobacterium stephanolepidis*, and *Mycobacterium salmoniphilum*, these species are reported as *M. chelonae* complex. The method also cannot distinguish between *M. mucogenicum* and *M. phocaicum*, so matches are reported as *M. mucogenicum/phocaicum* group. The MALDI-TOF MS results for isolates were used to adjust the presumptive NTM CFU/mL by multiplying the presumptive CFU/mL by the percentage of the isolates from the sample that were identified as NTM.

NTM qPCR

qPCR analysis followed the guidance of the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) and Environmental Microbiology Minimum Information Guidelines (EMMI) (92, 93). Gene copies of the *Mycobacterium atpE* gene were quantified using the Radomski et al. assay (FatpE: 5'-CGGYGCCGGTATCGGYGA-3'; RatpE: 5'-CGAAGACGAACARSGCCAT-3') (41), generating an approximately 164 base pair amplicon, with the modification that EvaGreen was used instead of probe chemistry (15). qPCR was performed using 96-well plates containing 10 µL reactions composed of 5 µL of master mix (Biotium FastEvaGreen 2x, final concentration 1x, cat. no. 31003), 0.5 µL each 10 µM forward and reverse primers (final concentration of 0.5 µM, Integrated DNA Technologies [IDT], Coralville, IA, USA), 0.25 mL of 25 mg/mL bovine serum albumin (Thermo Fisher Scientific, cat. no. AM2616, final concentration 0.625 mg/mL), 2.75 µL of ultrapure water and 1 µL of template. Samples were analyzed using a real-time PCR system (Applied Biosciences QuantStudio 3, Thermo Fisher Scientific) using the following cycling conditions: 95°C for 5 minutes, 35 cycles of 95°C for 20 seconds, 59.6°C for 30 seconds, and 72°C for 30 seconds. Melt curve analysis was performed for all plates. Standard curves were prepared using synthetic DNA (gBlock, IDT, USA), consisting of the amplicon with 30 base pairs of neutral adaptors on both ends. Standard curves were included on each 96-well plate. Negative controls with ultrapure water in place of template were also included on each 96-well plate. The LOD (41 gc/rx) and LOQ (41 gc/rxn) were

681 determined using serial dilutions of the standards. Samples were diluted to reduce inhibition, with
682 dilution factors ranging from undiluted to 1:110. Inhibition was assessed on the majority of
683 samples (88%, n=141) by analyzing samples at two dilution levels and comparing the actual Cq to
684 the theoretical Cq. Thresholds were automatically set by the qPCR instrument. All plates were
685 required to meet a minimum efficiency of 85% and R^2 standard curve results of at least 0.98.
686 Results were converted from gc/rxn to gc/L using the volume of template per reaction (1 µL), the
687 dilution factor, the DNA extraction elution volume (100 µL), and the filtered sample volume.
688
689 ***Metagenomic analysis***
690 The filter effluent, finished water, and full flush samples from distribution system Sites A, B, and
691 E from the December, March, June, and October sampling events (n=20) and three negative
692 controls (pooled filter controls, filtration controls, and DNA extraction controls) were submitted
693 for shotgun metagenomic sequencing. Library preparation, sequencing, and de-multiplexing were
694 performed by the University of Michigan Advanced Genomics Core. Paired-end sequence libraries
695 were prepared using the NEBNext® Ultra™ II FS DNA Library Prep Kit for Illumina (New
696 England BioLabs Inc., Ipswich, MA, USA, cat. no. E7805S). Sequencing was performed using the
697 NovaSeq 6000 system and an SP flow cell with 500 cycles, producing 250 nucleotide paired-end
698 reads.
699
700 Bioinformatic analysis of reads followed the pipeline described in Vosloo et al. (94). Briefly, pre-
701 processing was performed using fastp (v0.20.0) (95) to remove adaptors and perform initial quality
702 filtering. Reads mapping to the UniVec_Core database (ftp://ftp.ncbi.nlm.nih.gov/pub/UniVec/)
703 and negative controls were removed as potential contamination. The cleaned reads were pooled
704 and *de novo* co-assembly into contigs was performed using metaSPAdes (v.3.13.1) with kmer sizes
705 of 21, 33, 55, 77, 99, and 127 (96). Contigs less than 1 kilobase pair were removed using seqtk
706 (https://github.com/lh3/seqtk) and redundant contigs were removed using the *dedupe* function of
707 BBTools (v38.76) (97). The quality of the assembly was determined using QUAST (v.5.0.2) (98)
708 and assembly validation was performed by calculating the mapping rate of processed reads to
709 contigs. Binning was performed using Anvi'o (v6.1) workflow for the analysis and visualization
710 of omics data. Binning algorithms CONCOCT (v.1.1.1), MetaBAT (v.2.12.1), and MaxBin
711 (v.2.2.4) (99–102) were used together with the bin aggregation software DAS Tools (v.1.1.0) to

712 select the highest quality bins with the least redundancy (102). Bin statistics, including bin size,
713 GC content, and number of contigs, were determined using the summarize function in Anvi'o.
714 Further bin quality estimates, including completeness, redundancy, and strain heterogeneity, were
715 determined using CheckM (v 1.0.18) (104). To improve bin quality, bins with at least 50%
716 completeness were reassembled using metaSPAdes with the same kmer sizes previously used. Re-
717 binning was performed using the same binning strategy and manually curated using Anvi'o.
718 Duplicate bins were dereplicated using dRep (v2.6.2) and clustered into species-level
719 representative genomes using a 95% average nucleotide identity (105). The species-level
720 representative genomes were classified using the Genome Taxonomy Dataset Toolkit (GTDTk,
721 v0.3.2) (106) and final bin statistics and quality were determined as described previously. CoverM
722 (v0.4.0) was used to calculate coverage of the MAGs across samples with RPKM as a metric for
723 relative abundance (107). The covered_bases parameter of coverM was also used to calculate the
724 number of bases covered by one or more reads at a coverage threshold of 25%, indicating that only
725 bases with sequencing read coverage equal to or greater than 25% of the expected coverage depth
726 were considered as covered.
727
728 ***Data analysis***
729 Data cleaning and analyses were performed using R (version 4.1.1) and R Studio (1.4.1717) (108,
730 109). R packages used for analyses include *ggplot, dplyr, lubridate, viridis, readxl, and stats*. For
731 the qPCR data analysis and plotting, values less than the LOQ were set at one-half the LOQ. For
732 physicochemical parameters, values less than the LOD were set at one-half the LOD. Statistical
733 analyses were performed using the *stats* package. Differences between sample locations, sample
734 types, and by season were calculated using the Wilcoxon signed-rank test using a significance
735 threshold of 0.05. Correlation analysis and correlation plot were performed using the functions
736 "cor" and "cor.test" from the R package *stats* and the R package *corrplot* using Kendall's tau b.
737
738 **Acknowledgments**
739 The authors gratefully acknowledge the staff of the City of Ann Arbor WTP and building owners
740 and staff for their support of this work. The authors thank Ruqaiya Siddiqui, Michael Bachman,
741 Stephanie Agozino, and Jennifer Furlong for supporting MALDI-TOF MS analysis and Bridget
742 Hegarty for helpful discussions. The authors thank Ameet Pinto and Solize Vosloo for supporting

743 the analysis of the metagenomics data. Funding was provided by the Blue Sky Initiative (College
744 of Engineering, University of Michigan) and Water Research Foundation Project Number 4721.
745 K.S.D was supported by a National Science Foundation Graduate Research Fellowship (grant
746 number DGE-1256260) and a University of Michigan Rackham Predoctoral Fellowship. This
747 research was supported by work performed by The University of Michigan Microbiome Core.
748

749 **Data Availability Statement**

750 The raw sequencing data are available through the National Center for Biotechnology Information
751 (NCBI) under BioProject No. PRJNA1081894.
752

753 **Supplemental Information**

754 Supplementary materials are available in the Supplementary Information file.

References

1. Park SC, Kang MJ, Han CH, Lee SM, Kim CJ, Lee JM, Kang YA. 2019. Prevalence, incidence, and mortality of nontuberculous mycobacterial infection in Korea: a nationwide population-based study. BMC Pulmonary Medicine 19.

2. Shah NM, Davidson JA, Anderson LF, Lalor MK, Kim J, Thomas HL, Lipman M, Abubakar I. 2016. Pulmonary *Mycobacterium avium-intracellulare* is the main driver of the rise in non-tuberculous mycobacteria incidence in England, Wales and Northern Ireland, 2007–2012. BMC Infectious Diseases 16:195.

3. Prevots DR, Loddenkemper R, Sotgiu G, Migliori GB. 2017. Nontuberculous mycobacterial pulmonary disease: an increasing burden with substantial costs. European Respiratory Journal 49:1700374.

4. Collier SA, Deng L, Adam EA, Benedict KM, Beshearse EM, Blackstock AJ, Bruce BB, Derado G, Edens C, Fullerton KE, Gargano JW, Geissler AJ, Havelaar AH, Hill VR, Hoekstra RM, Reddy SC, Scallan E, Yoder JS, Beach MJ. 2021. Estimate of burden and direct healthcare cost of infectious waterborne disease in the United States. Emerging Infectious Diseases 27.

5. Greco SL, Drudge C, Fernandes R, Kim J, Copes R. 2020. Estimates of healthcare utilisation and deaths from waterborne pathogen exposure in Ontario, Canada. Epidemiology and Infection 148.

6. Gerdes ME, Miko S, Kunz JM, Hannapel EJ, Hlavsa MC, Hughes MJ, Stuckey MJ, Francois Watkins LK, Cope JR, Yoder JS, Hill VR, Collier SA. 2023. Estimating Waterborne Infectious Disease Burden by Exposure Route, United States, 2014. Emerg Infect Dis 29.

7. Dowdell K, Haig S-J, Caverly LJ, Shen Y, LiPuma JJ, Raskin L. 2019. Nontuberculous mycobacteria in drinking water systems – the challenges of characterization and risk mitigation. Current Opinion in Biotechnology 57:127–136.

8. Nishiuchi Y, Iwamoto T, Maruyama F. 2017. Infection sources of a common non-tuberculous mycobacterial pathogen, *Mycobacterium avium* complex. Frontiers in Medicine 4:27.

9. Lande L. 2019. Environmental niches for NTM and their impact on NTM disease, p. 131–144. *In* Griffith, DE (ed.), Nontuberculous Mycobacterial Disease: A Comprehensive Approach to Diagnosis and Management. Humana Press.

10. Hamilton KA, Weir MH, Haas CN. 2017. Dose response models and a quantitative microbial risk assessment framework for the *Mycobacterium avium* complex that account for recent developments in molecular biology, taxonomy, and epidemiology. Water Research 109:310–326.

11. Proctor C, Garner E, Hamilton KA, Ashbolt NJ, Caverly LJ, Falkinham JO, Haas CN, Prevost M, Prevots DR, Pruden A, Raskin L, Stout J, Haig S-J. 2022. Tenets of a holistic approach to drinking water-associated pathogen research, management, and communication. Water Research 211:117997.

12. Ley CJ, Proctor CR, Singh G, Ra K, Noh Y, Odimayomi T, Salehi M, Julien R, Mitchell J, Nejadhashemi AP, Whelton AJ, Aw TG. 2020. Drinking water microbiology in a water-efficient building: stagnation, seasonality, and physicochemical effects on opportunistic pathogen and total bacteria proliferation. Environ Sci: Water Res Technol 6:2902–2913.

798 13. Logan-Jackson AR, Flood M, Rose JB. 2021. Enumeration and characterization of five
799 pathogenic *Legionella* species from large research and educational buildings. Environ Sci:
800 Water Res Technol 7:321–334.

801 14. Haig S-J, Kotlarz N, LiPuma JJ, Raskin L. 2018. A high-throughput approach for
802 identification of nontuberculous mycobacteria in drinking water reveals relationship
803 between water age and *Mycobacterium avium*. mBio 9:e02354-17.

804 15. Haig S-J, Kotlarz N, Kalikin LM, Chen T, Guikema S, LiPuma JJ, Raskin L. 2020.
805 Emerging investigator series: bacterial opportunistic pathogen gene markers in municipal
806 drinking water are associated with distribution system and household plumbing
807 characteristics. Environmental Science: Water Research & Technology
808 https://doi.org/10.1039/D0EW00723D.

809 16. Hozalski RM, LaPara TM, Zhao X, Kim T, Waak MB, Burch T, McCarty M. 2020. Flushing
810 of Stagnant Premise Water Systems after the COVID-19 Shutdown Can Reduce Infection
811 Risk by *Legionella* and *Mycobacterium* spp. Environ Sci Technol 54:15914–15924.

812 17. Lu J, Buse H, Struewing I, Zhao A, Lytle D, Ashbolt N. 2017. Annual variations and effects
813 of temperature on *Legionella* spp. and other potential opportunistic pathogens in a
814 bathroom. Environmental Science and Pollution Research 24:2326–2336.

815 18. Aw TG, Scott L, Jordan K, Ra K, Ley C, Whelton AJ. 2022. Prevalence of opportunistic
816 pathogens in a school building plumbing during periods of low water use and a transition to
817 normal use. International Journal of Hygiene and Environmental Health 241:113945.

19. Zhang C, Struewing I, Mistry JH, Wahman DG, Pressman J, Lu J. 2021. *Legionella* and other opportunistic pathogens in full-scale chloraminated municipal drinking water distribution systems. Water Research 205:117571.

20. Rahmatika I, Simazaki D, Kurisu F, Furumai H, Kasuga I. 2023. Occurrence and diversity of nontuberculous mycobacteria affected by water stagnation in building plumbing. Water Supply 23:5017–5028.

21. Pfaller S, King D, Mistry JH, Alexander M, Abulikemu G, Pressman JG, Wahman DG, Donohue MJ. 2021. Chloramine Concentrations within Distribution Systems and Their Effect on Heterotrophic Bacteria, Mycobacterial Species, and Disinfection Byproducts. Water Research 117689.

22. Pfaller S, King D, Mistry JH, Donohue M. 2022. Occurrence revisited: *Mycobacterium avium* and *Mycobacterium intracellulare* in potable water in the USA. Appl Microbiol Biotechnol 106:2715–2727.

23. King DN, Donohue MJ, Vesper SJ, Villegas EN, Ware MW, Vogel ME, Furlong EF, Kolpin DW, Glassmeyer ST, Pfaller S. 2016. Microbial pathogens in source and treated waters from drinking water treatment plants in the United States and implications for human health. Science of the Total Environment 562:987–995.

24. Li Q, Yu S, Li L, Liu G, Gu Z, Liu M, Liu Z, Ye Y, Xia Q, Ren L. 2017. Microbial communities shaped by treatment processes in a drinking water treatment plant and their contribution and threat to drinking water safety. Frontiers in Microbiology 8:2465.

838 25. Le Dantec C, Duguet JP, Montiel A, Dumoutier N, Dubrou S, Vincent V. 2002. Occurrence
839 of mycobacteria in water treatment lines and in water distribution systems. Applied and
840 Environmental Microbiology 68:5318–5325.

841 26. Thomas V, Loret J-F, Jousset M, Greub G. 2008. Biodiversity of amoebae and amoebae-
842 resisting bacteria in a drinking water treatment plant. Environmental Microbiology
843 10:2728–2745.

844 27. Kotlarz N, Rockey N, Olson TM, Haig S-J, Sanford L, LiPuma JJ, Raskin L. 2018. Biofilms
845 in full-scale drinking water ozone contactors contribute viable bacteria to ozonated water.
846 Environmental Science & Technology 52:2618–2628.

847 28. Donohue MJ, Vesper S, Mistry J, Donohue JM. 2019. Impact of Chlorine and Chloramine
848 on the Detection and Quantification of *Legionella pneumophila* and *Mycobacterium*
849 Species. Applied and Environmental Microbiology 85.

850 29. Pryor M, Springthorpe S, Riffard S, Brooks T, Huo Y, Davis G, Sattar SA. 2004.
851 Investigation of opportunistic pathogens in municipal drinking water under different supply
852 and treatment regimes. Water Science and Technology 50:83–90.

853 30. Donohue MJ, Mistry JH, Donohue JM, O'Connell K, King D, Byran J, Covert T, Pfaller S.
854 2015. Increased frequency of nontuberculous mycobacteria detection at potable water taps
855 within the United States. Environmental Science and Technology 49:6127–6133.

856 31. Proctor CR, Rhoads WJ, Keane T, Salehi M, Hamilton K, Pieper KJ, Cwiertny DM, Prévost
857 M, Whelton AJ. 2020. Considerations for large building water quality after extended
858 stagnation. AWWA Water Science 2:e1186.

859 32. Dowdell KS, Healy HG, Joshi S, Grimard-Conea M, Pitell S, Song Y, Ley C, Kennedy LC,
860 Vosloo S, Huo L, Haig S-J, Hamilton KA, Nelson KL, Pinto A, Prévost M, Proctor CR,
861 Raskin L, Whelton AJ, Garner E, Pieper KJ, Rhoads WJ. 2023. *Legionella pneumophila*
862 occurrence in reduced-occupancy buildings in 11 cities during the COVID-19 pandemic.
863 Environ Sci: Water Res Technol 10.1039.D3EW00278K.

864 33. Richard R, Boyer TH. 2021. Pre- and post-flushing of three schools in Arizona due to
865 COVID-19 shutdown. AWWA Water Science 3.

866 34. Rhoads WJ, Hammes F. 2021. Growth of *Legionella* during COVID-19 lockdown
867 stagnation. Environmental Science: Water Research & Technology
868 https://doi.org/10.1039/D0EW00819B.

869 35. Thomson R, Carter R, Gilpin C, Coulter C, Hargreaves M. 2008. Comparison of Methods
870 for Processing Drinking Water Samples for the Isolation of *Mycobacterium avium* and
871 *Mycobacterium intracellulare*. Applied and Environmental Microbiology 74:3094–3098.

872 36. Williams MD, Falkinham JO. 2018. Effect of Cetylpyridinium Chloride (CPC) on Colony
873 Formation of Common Nontuberculous Mycobacteria. Pathogens 7:79.

874 37. Fernandes HMZ, Conceição EC, Gomes KM, da Silva MG, Dias RCS, Duarte RS. 2019.
875 Recovery of Non-tuberculous Mycobacteria from Water is Influenced by Phenotypic
876 Characteristics and Decontamination Methods. Current Microbiology
877 https://doi.org/10.1007/s00284-019-01704-w.

878 38. Radomski N, Cambau E, Moulin L, Haenn S, Moilleron R, Lucas FS. 2010. Comparison of
879 Culture Methods for Isolation of Nontuberculous Mycobacteria from Surface Waters.
880 Applied and Environmental Microbiology 76:3514–3520.

881 39. Wang H, Bédard E, Prévost M, Camper AK, Hill VR, Pruden A. 2017. Methodological
882 approaches for monitoring opportunistic pathogens in premise plumbing: A review. Water
883 Research 117:68–86.

884 40. Aw TG, Rose JB. 2012. Detection of pathogens in water: from phylochips to qPCR to
885 pyrosequencing. Current Opinion in Biotechnology 23:422–430.

886 41. Radomski N, Roguet A, Lucas FS, Veyrier FJ, Cambau E, Accrombessi H, Moilleron R,
887 Behr MA, Moulin L. 2013. *atpE* gene as a new useful specific molecular target to quantify
888 *Mycobacterium* in environmental samples. BMC Microbiology 13:277.

889 42. Chern EC, King D, Haugland R, Pfaller S. 2015. Evaluation of quantitative polymerase
890 chain reaction assays targeting *Mycobacterium avium*, *M. intracellulare*, and *M. avium*
891 subspecies *paratuberculosis* in drinking water biofilms. Journal of Water and Health
892 13:131–139.

893 43. Dowdell KS, Raskin L, Olson T, Haig S-J, Dai D, Edwards M, Pruden A. 2022. Methods
894 for Detecting and Differentiating Opportunistic Premise Plumbing Pathogens (OPPPs) to
895 Determine Efficacy of Control and Treatment Technologies. Report 4721. Water Research
896 Foundation, Denver, CO.

897 44. Lee E-S, Lee M-H, Kim B-S. 2015. Evaluation of propidium monoazide-quantitative PCR
898 to detect viable *Mycobacterium fortuitum* after chlorine, ozone, and ultraviolet disinfection.
899 International Journal of Food Microbiology 210:143–148.

900 45. Ditommaso S, Giacomuzzi M, Memoli G, Cavallo R, Curtoni A, Avolio M, Silvestre C,
901 Zotti CM. 2019. Reduction of turnaround time for non-tuberculous mycobacteria detection
902 in heater-cooler units by propidium monoazide real-time PCR. Journal of Hospital Infection
903 https://doi.org/10.1016/j.jhin.2019.10.010.

904 46. Nocker A, Sossa KE, Camper AK. 2007. Molecular monitoring of disinfection efficacy
905 using propidium monoazide in combination with quantitative PCR. Journal of
906 Microbiological Methods 70:252–260.

907 47. Gebert MJ, Delgado-Baquerizo M, Oliverio AM, Webster TM, Nichols LM, Honda JR,
908 Chan ED, Adjemian J, Dunn RR, Fierer N. 2018. Ecological analyses of mycobacteria in
909 showerhead biofilms and their relevance to human health. mBio 9:1–15.

910 48. Walsh CM, Gebert MJ, Delgado-Baquerizo M, Maestre FT, Fierer N. 2019. A Global
911 Survey of Mycobacterial Diversity in Soil. Appl Environ Microbiol 85:e01180-19.

912 49. Honda JR, Hasan NA, Davidson RM, Williams MD, Epperson LE, Reynolds PR, Smith T,
913 Iakhiaeva E, Bankowski MJ, Wallace Jr. RJ, Chan ED, Falkinham III JO, Strong M, Wallace
914 RJ, Chan ED, Falkinham JO, Strong M. 2016. Environmental nontuberculous mycobacteria
915 in the Hawaiian Islands. PLoS Neglected Tropical Diseases 10:e0005068.

916 50. Pinto AJ, Raskin L. 2012. PCR Biases Distort Bacterial and Archaeal Community Structure
917 in Pyrosequencing Datasets. PLoS ONE 7:e43093.

51. Bonk F, Popp D, Harms H, Centler F. 2018. PCR-based quantification of taxa-specific abundances in microbial communities: Quantifying and avoiding common pitfalls. Journal of Microbiological Methods 153:139–147.

52. Crossette E, Gumm J, Langenfeld K, Raskin L, Duhaime M, Wigginton K. 2021. Metagenomic Quantification of Genes with Internal Standards. mBio 12:e03173-20.

53. Nayfach S, Pollard KS. 2016. Toward Accurate and Quantitative Comparative Metagenomics. Cell 166:1103–1116.

54. Caverly LJ, Carmody LA, Haig SJ, Kotlarz N, Kalikin LM, Raskin L, LiPuma JJ. 2016. Culture-independent identification of nontuberculous mycobacteria in cystic fibrosis respiratory samples. PLoS ONE 11:e0153876.

55. Johansen MD, Herrmann J-L, Kremer L. 2020. Non-tuberculous mycobacteria and the rise of *Mycobacterium abscessus*. Nature Reviews Microbiology https://doi.org/10.1038/s41579-020-0331-1.

56. Adjemian J, Olivier KN, Prevots DR. 2018. Epidemiology of pulmonary nontuberculous mycobacterial sputum positivity in patients with cystic fibrosis in the United States, 2010-2014. Annals of the American Thoracic Society 15:817–825.

57. Donohue MJ. 2021. Epidemiological risk factors and the geographical distribution of eight Mycobacterium species. BMC Infect Dis 21:258.

936 58. Cazals M, Bédard E, Faucher SP, Prévost M. 2023. Factors Affecting the Dynamics of
937 *Legionella pneumophila* , Nontuberculous Mycobacteria, and Their Host *Vermamoeba*
938 *vermiformis* in Premise Plumbing. ACS EST Water 3:3874–3883.

939 59. Delafont V, Mougari F, Cambau E, Joyeux M, Bouchon D, Héchard Y, Moulin L. 2014.
940 First evidence of amoebae-mycobacteria association in drinking water network.
941 Environmental Science and Technology 48:11872–11882.

942 60. Spencer-Williams I, Meyer M, DePas W, Elliott E, Haig S-J. 2023. Assessing the Impacts
943 of Lead Corrosion Control on the Microbial Ecology and Abundance of Drinking-Water-
944 Associated Pathogens in a Full-Scale Drinking Water Distribution System. Environ Sci
945 Technol acs.est.3c05272.

946 61. Wang H, Masters S, Falkinham JO, Edwards MA, Pruden A. 2015. Distribution system
947 water quality affects responses of opportunistic pathogen gene markers in household water
948 heaters. Environmental Science & Technology 49:8416–8424.

949 62. Gomez-Alvarez V, Ryu H, Tang M, McNeely M, Muhlen C, Urbanic M, Williams D, Lytle
950 D, Boczek L. 2023. Assessing residential activity in a home plumbing system simulator:
951 monitoring the occurrence and relationship of major opportunistic pathogens and
952 phagocytic amoebas. Front Microbiol 14:1260460.

953 63. Ringuet H, Akoua-Koffi C, Honore S, Varnerot A, Vincent V, Berche P, Gaillard JL, Pierre-
954 Audigier C. 1999. *hsp65* sequencing for identification of rapidly growing mycobacteria.
955 Journal of Clinical Microbiology 37:852–857.

956 64. Lee H, Park HJ, Cho SN, Bai GH, Kim SJ. 2000. Species identification of mycobacteria by
957 PCR-restriction fragment length polymorphism of the *rpoB* gene. Journal of Clinical
958 Microbiology 38:2966–2971.

959 65. Adékambi T, Drancourt M. 2004. Dissection of phylogenetic relationships among 19 rapidly
960 growing *Mycobacterium* species by 16S rRNA, *hsp65*, *sodA*, *recA* and *rpoB* gene
961 sequencing. International Journal of Systematic and Evolutionary Microbiology 54:2095–
962 2105.

963 66. Adékambi T, Colson P, Drancourt M. 2003. *rpoB*-Based Identification of Nonpigmented
964 and Late-Pigmenting Rapidly Growing Mycobacteria. Journal of Clinical Microbiology
965 41:5699–5708.

966 67. Alcaide F, Amlerová J, Bou G, Ceyssens PJ, Coll P, Corcoran D, Fangous M-S, González-
967 Álvarez I, Gorton R, Greub G, Hery-Arnaud G, Hrábak J, Ingebretsen A, Lucey B,
968 Marekovič I, Mediavilla-Gradolph C, Monté MR, O'Connor J, O'Mahony J, Opota O,
969 O'Reilly B, Orth-Höller D, Oviaño M, Palacios JJ, Palop B, Pranada AB, Quiroga L,
970 Rodríguez-Temporal D, Ruiz-Serrano MJ, Tudó G, Van den Bossche A, van Ingen J,
971 Rodriguez-Sanchez B. 2018. How to: identify non-tuberculous *Mycobacterium* species
972 using MALDI-TOF mass spectrometry. Clinical Microbiology and Infection 24:599–603.

973 68. Rodriguez-Temporal D, Rodríguez-Sánchez B, Alcaide F. 2020. Evaluation of MALDI
974 Biotyper Interpretation Criteria for Accurate Identification of Nontuberculous
975 Mycobacteria. J Clin Microbiol 58.

976 69. Rindi L, Puglisi V, Franconi I, Fais R, Lupetti A. 2022. Rapid and Accurate Identification
977 of Nontuberculous Mycobacteria Directly from Positive Primary MGIT Cultures by
978 MALDI-TOF MS. Microorganisms 10:1447.

979 70. El Khéchine A, Couderc C, Flaudrops C, Raoult D, Drancourt M. 2011. Matrix-Assisted
980 Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Identification of
981 Mycobacteria in Routine Clinical Practice. PLoS ONE 6:e24720.

982 71. van der Wielen PWJJ, Heijnen L, van der Kooij D. 2013. Pyrosequence Analysis of the
983 *hsp65* Genes of Nontuberculous Mycobacterium Communities in Unchlorinated Drinking
984 Water in the Netherlands. Applied and Environmental Microbiology 79:6160–6166.

985 72. Waak MB, LaPara TM, Hallé C, Hozalski RM. 2019. Nontuberculous Mycobacteria in Two
986 Drinking Water Distribution Systems and the Role of Residual Disinfection. Environ Sci
987 Technol 53:8563–8573.

988 73. Hull NM, Holinger EP, Ross KA, Robertson CE, Harris JK, Stevens MJ, Pace NR. 2017.
989 Longitudinal and source-to-tap New Orleans, LA, U.S.A. drinking water microbiology.
990 Environmental Science and Technology 51:4220–4229.

991 74. Genilloud O. 2015. *Micromonospora*, p. 1–28. *In* Whitman, WB, Rainey, F, Kämpfer, P,
992 Trujillo, M, Chun, J, DeVos, P, Hedlund, B, Dedysh, S (eds.), Bergey's Manual of
993 Systematics of Archaea and Bacteria, 1st ed. Wiley.

994 75. Logan NA, Vos PD. 2015. *Bacillus*, p. 1–163. *In* Whitman, WB, Rainey, F, Kämpfer, P,
995 Trujillo, M, Chun, J, DeVos, P, Hedlund, B, Dedysh, S (eds.), Bergey's Manual of
996 Systematics of Archaea and Bacteria, 1st ed. Wiley.

997 76. Logan NA, Vos PD. 2015. *Brevibacillus*, p. 1–22. *In* Whitman, WB, Rainey, F, Kämpfer,
998 P, Trujillo, M, Chun, J, DeVos, P, Hedlund, B, Dedysh, S (eds.), Bergey's Manual of
999 Systematics of Archaea and Bacteria, 1st ed. Wiley.

1000 77. Priest FG. 2015. *Paenibacillus*, p. 1–40. *In* Whitman, WB, Rainey, F, Kämpfer, P, Trujillo,
1001 M, Chun, J, DeVos, P, Hedlund, B, Dedysh, S (eds.), Bergey's Manual of Systematics of
1002 Archaea and Bacteria, 1st ed. Wiley.

1003 78. Grigg C, Jackson KA, Barter D, Czaja CA, Johnston H, Lynfield R, Snippes Vagnone P,
1004 Tourdot L, Spina N, Dumyati G, Cassidy PM, Pierce R, Henkle E, Prevots DR, Salfinger
1005 M, Winthrop KL, Charles Toney N, Magill SS. 2023. Epidemiology of Pulmonary and
1006 Extrapulmonary Nontuberculous Mycobacteria Infections at 4 US Emerging Infections
1007 Program Sites: A 6-Month Pilot. Clinical Infectious Diseases ciad214.

1008 79. Spaulding AB, Lai YL, Zelazny AM, Olivier KN, Kadri SS, Rebecca Prevots D, Adjemian
1009 J. 2017. Geographic distribution of nontuberculous mycobacterial species identified among
1010 clinical isolates in the United States, 2009-2013. Annals of the American Thoracic Society
1011 14:1655–1661.

1012 80. Liang J, Swanson CS, Wang L, He Q. 2021. Impact of building closures during the COVID-
1013 19 pandemic on *Legionella* infection risks. American Journal of Infection Control
1014 S0196655321005952.

1015 81. Huang J, Chen S, Ma X, Yu P, Zuo P, Shi B, Wang H, Alvarez PJJ. 2021. Opportunistic
1016 pathogens and their health risk in four full-scale drinking water treatment and distribution
1017 systems. Ecological Engineering 160:106134.

1018 82. Pinto AJ, Xi C, Raskin L. 2012. Bacterial community structure in the drinking water
1019 microbiome is governed by filtration processes. Environmental Science and Technology
1020 46:8851–8859.

1021 83. Griffith DE, Aksamit T, Brown-Elliott BA, Catanzaro A, Daley C, Gordin F, Holland SM,
1022 Horsburgh R, Huitt G, Iademarco MF, Iseman M, Olivier K, Ruoss S, Von Reyn CF,
1023 Wallace RJ, Winthrop K. 2007. An official ATS/IDSA statement: Diagnosis, treatment, and
1024 prevention of nontuberculous mycobacterial diseases 175:367–416.

1025 84. Moore JE, Kruijshaar ME, Ormerod LP, Drobniewski F, Abubakar I. 2010. Increasing
1026 reports of non-tuberculous mycobacteria in England, Wales and Northern Ireland, 1995-
1027 2006. BMC Public Health 10:612.

1028 85. Jones MM, Winthrop KL, Nelson SD, Duvall SL, Patterson OV, Nechodom KE, Findley
1029 KE, Radonovich LJ, Samore MH, Fennelly KP. 2018. Epidemiology of nontuberculous
1030 mycobacterial infections in the U.S. Veterans Health Administration. PLoS ONE
1031 13:e0197976.

1032 86. Alexander KJ, Furlong JL, Baron JL, Rihs JD, Stephenson D, Perry JD, Stout JE. 2021.
1033 Evaluation of a new culture medium for isolation of nontuberculous mycobacteria from
1034 environmental water samples. PLoS ONE 16:e0247166.

1035 87. Ditommaso S, Giacomuzzi M, Memoli G, Garlasco J, Curtoni A, Iannaccone M, Zotti CM.
1036 2022. Chemical susceptibility testing of non-tuberculous mycobacterium strains and other
1037 aquatic bacteria: Results of a study for the development of a more sensitive and simple

method for the detection of NTM in environmental samples. Journal of Microbiological Methods 193:106405.

88. van der Wielen PWJJ, van der Kooij D. 2013. Nontuberculous Mycobacteria, Fungi, and Opportunistic Pathogens in Unchlorinated Drinking Water in the Netherlands. Appl Environ Microbiol 79:825–834.

89. Rossman L, Woo H, Tryby M, Shang F, Janke R, Haxton T. 2020. EPANET (2.2.0). U.S. Environmental Protection Agency, Washington, DC, USA.

90. Vosloo S, Sevillano M, Pinto A. 2019. Modified DNeasy PowerWater Kit® protocol for DNA extractions from drinking water samples.

91. Standard Methods Committee of the American Public Health Association, American Water Works Association, and Water Environment Federation. 2017. 9215 Heterotrophic Plate Count, p. . *In* Lipps, WC, Baxter, TE, Braun-Howland, E (eds.), Standard Methods For the Examination of Water and Wastewater, 23rd ed. APHA Press, Washington, D.C.

92. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clinical Chemistry 55:611–622.

93. Borchardt MA, Boehm AB, Salit M, Spencer SK, Wigginton KR, Noble RT. 2021. The Environmental Microbiology Minimum Information (EMMI) Guidelines: qPCR and dPCR Quality and Reporting for Environmental Microbiology. Environ Sci Technol acs.est.1c01767.

1059 94. Vosloo S, Huo L, Anderson CL, Dai Z, Sevillano M, Pinto A. 2021. Evaluating de Novo
1060 Assembly and Binning Strategies for Time Series Drinking Water Metagenomes.
1061 Microbiology Spectrum 9:18.

1062 95. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor.
1063 Bioinformatics 34:i884–i890.

1064 96. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile
1065 metagenomic assembler. Genome Res 27:824–834.

1066 97. Bushnell B. 2018. BBTools: a suite of fast, multithreaded bioinformatics tools designed for
1067 analysis of DNA and RNA sequence data. Joint Genome Institute.

1068 98. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome
1069 assemblies 4.

1070 99. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015.
1071 Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319.

1072 100. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ,
1073 Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and
1074 composition. Nat Methods 11:1144–1146.

1075 101. Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome
1076 reconstruction from metagenome assemblies 13.

1077 102. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to
1078 recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607.

1079 103. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018.
1080 Recovery of genomes from metagenomes via a dereplication, aggregation and scoring
1081 strategy. Nat Microbiol 3:836–843.

1082 104. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing
1083 the quality of microbial genomes recovered from isolates, single cells, and metagenomes.
1084 Genome Res 25:1043–1055.

1085 105. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate
1086 genomic comparisons that enables improved genome recovery from metagenomes through
1087 de-replication. ISME J 11:2864–2868.

1088 106. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify
1089 genomes with the Genome Taxonomy Database. Bioinformatics btz848.

1090 107. Aroney STN, Newell RJP, Nissen J, Camargo AP, Tyson GW, Woodcroft BJ. 2024.
1091 CoverM: Read coverage calculator for metagenomics (0.7.0).

1092 108. R Core Team. 2021. R: A language and environment for statistical computing. R Foundation
1093 for Statistical Computing, Vienna, Austria.

1094 109. RStudio Team. 2020. RStudio: Integrated Development for R. RStudio, Boston, MA.

1095