NONTUBERCULOUS MYCOBACTERIA: PREVALENCE IN NEW JERSEY PRIVATE WELL BIOFILMS TODAY, POTENTIAL FOR CHANGING RISK TOMORROW

By

SOPHIA MARIE BLANC

A thesis submitted to the

School of Graduate Studies

Rutgers, the State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Civil and Environmental Engineering

Written under the direction of

Nicole Fahrenfeld

And approved by

____________________________

____________________________

____________________________

New Brunswick, New Jersey

May 2021
ABSTRACT OF THE THESIS

Nontuberculous Mycobacteria in New Jersey Private Well Biofilms Today,

Potential for Changing Risk Tomorrow

Thesis Director:

Nicole Fahrenfeld

Household plumbing biofilms can harbor and transmit bacterial pathogens. Pulmonary infections by nontuberculous mycobacteria (NTM) can occur from this transmission route. NTM infections are increasing around the world and in New Jersey (NJ). To understand the abundance of NTM in NJ private wells today, a field study was performed (Ch. II). To evaluate how NTM abundance may change in water environments in the future, a critical literature review was performed (Ch. III). For the field study, plumbing biofilm samples were collected from sinks and showerheads in homes using private wells (N=19) and in homes of NTM patients (N=5). DNA extracts were analyzed by qPCR to quantify mycobacterial marker genes and by amplicon sequencing to describe the microbiomes where NTM were observed. Water samples were analyzed for basic water quality parameters and fecal indicator organisms. Participants completed surveys about their wells and home water systems to enable testing of potential relationships between these environmental factors and the microbial communities. NTM were observed in more than half of private well biofilm samples using qPCR and in all of the selected samples analyzed by amplicon sequencing (N=29/70), even when below detection by qPCR. Samples from patient homes, most of whom used public water supply, had similar abundances of NTM as samples from private wells. Physiography and within-home location (e.g., kitchen sink) explained some variation in concentrations of mycobacteria genes. One microbial family with no known human pathogens, Rikenellaceae, was identified as a potential antagonist to mycobacteria using
linear discriminant analysis effect size (LEfSe). This study illuminated the widespread nature of NTM in private well water systems without disinfection selection pressures, and worked towards understanding ecological interactions that may aid or slow the growth of NTM toward ecological engineering of healthy plumbing microbiomes.

The critical literature review was performed systematically by searching specific terms in several online databases. Connections were made between NTM fate and transport, climate change, engineering decisions, and societal changes, and uncertainties highlighted. Environmental conditions discussed with respect to NTM risk included changing temperature, humidity, salinity, rainfall, and extreme weather events. NTM risk was then considered under climate/societal scenarios described by Intergovernmental Panel on Climate Change (IPCC) scientists. Findings indicate that the resilience of NTM under a variety of environmental conditions (e.g., warm temperatures, eutrophication) may increase their net prevalence in water environments under climate change, increasing exposure. Water management decisions may also influence exposure to NTM as water scarcity requires increased reliance on reclaimed water. Water managers may control risk of exposure through innovative water treatment processes and equitable water management decisions, turning towards an integrated One Water approach to reduce and/or mitigate the impacts of de facto reuse.
ACKNOWLEDGEMENT

This thesis would not be possible without the support of many individuals, primarily my advisor Dr. Nicole Fahrenfeld who supported me every step of the way. I cannot thank her enough for her technical, professional, and emotional guidance. Thanks to my committee, Dr. Jie Gong and Dr. Monica Mazurek for their willingness to offer insight about this work. Thanks to our partners at the Public Health Institute of New Jersey, Dr. Christopher Vinnard and Dr. Maria Gennaro, for their contributions to our study design, patient recruitment, and our final report. Thanks to my lab mates: to Dr. Alessia Eramo for providing my initial training so many years ago, to William Morales-Medina for his absolute willingness to help with anything he can, to Aishwarya Deshpande for her work in getting this project started, to Dion Pender, Cassidy Schneider, and Rachel Feldstein for playing crucial roles in our sampling efforts, to Melissa Tu for her dependable technical assistance, and to Ian Roback and Sasha Monti for their assistance with literature review. Thanks to Dr. David Robinson for his assistance with understanding the climate science. Thanks to Dr. Karen O’Neill for her collaboration about the social implications of this work. Thanks to Dr. Carrie Ferraro, Jeanne Herb, and the others with the Coastal Climate Risk and Resiliency (C2R2) Traineeship for their endless support throughout this process, and for their input regarding the systems thinking involved in this project. Thanks to the National Science Foundation C2R2 (Grant #1633557), the USGS/ NJWRRI Water Resources Research Act Program National Competitive Grants Program (#G20AS00026), the New Jersey Department of Health, and the Rutgers Sustainable Raritan River Group for their funding support. Finally, thanks to my family and friends for their support. Please note, Chapter II is comprised of an article that is in revision. The citation is below, and the author attribution is in Section VII.

# TABLE OF CONTENTS

ABSTRACT OF THE THESIS ........................................................................................................... ii

ACKNOWLEDGEMENT .............................................................................................................. iv

TABLE OF CONTENTS ............................................................................................................. v

LIST OF TABLES ....................................................................................................................... vii

LIST OF ILLUSTRATIONS ......................................................................................................... viii

I. INTRODUCTION .................................................................................................................... 1

1. Background ......................................................................................................................... 1

2. Gaps in understanding and thesis objectives .................................................................... 5

II. NONTUBERCULOUS MYCOBACTERIA IN THE BIOFILM MICROBIOME OF PRIVATE WELL AND PREMISE PLUMBING .............................................................................................................. 9

1. Background and introduction ............................................................................................ 9

2. Material & methods .......................................................................................................... 12

2.1. Participant recruitment .................................................................................................. 12

2.2. Biofilm collection ......................................................................................................... 13

2.3. Water sample collection & quality analyses ................................................................. 14

2.4. Biomolecular analyses for mycobacteria and microbial community ......................... 15

2.5. Statistical analyses ....................................................................................................... 18

3. Results & discussion ......................................................................................................... 21

3.1. Water quality ................................................................................................................ 21

3.2. Survey results ............................................................................................................... 23

3.3. qPCR for mycobacteria and total bacterial community ................................................ 24

3.4. Biofilm microbiome ..................................................................................................... 30

4. Conclusion ......................................................................................................................... 38

5. Acknowledgements ......................................................................................................... 39

6. Author attributions & citation ......................................................................................... 39

III. POTENTIAL FOR NONTUBERCULOUS MYCOBACTERIA PROLIFERATION IN NATURAL AND ENGINEERED WATER SYSTEMS DUE TO CLIMATE CHANGE: A CRITICAL LITERATURE REVIEW ......................................................................................................................... 40

1. Introduction ....................................................................................................................... 40

1.1. Disease prevalence and impact .................................................................................... 40

1.2. The nexus of pathogen, climate, and human interactions- a roadmap to risk analysis .. 41

2. Methods .......................................................................................................................... 43

3. Results and Discussion ................................................................................................. 43
3.1. Biological, ecological, and physiological characteristics ........................................43
3.2. Climate change considerations ..............................................................................47
3.3. Changes in engineering and water management practices ...................................53
3.4. Related societal dynamics ....................................................................................58
3.5. Potential scenarios for consideration .................................................................60
3.6. Uncertainties and recommendations for future work ........................................62
4. Conclusions ............................................................................................................65
IV. CONCLUSION .......................................................................................................66
V. APPENDIX: Ch II Supplementary Materials .........................................................68
VI. REFERENCES .......................................................................................................80
VII. ACKNOWLEDGEMENT OF PREVIOUS PUBLICATIONS .................................99
LIST OF TABLES

TABLE 1 Water quality results ................................................................. 21

TABLE 2 NTM private well biofilm data ....................................................... 24

TABLE 3 NTM pulmonary patient biofilm data ............................................. 25

TABLE 4 Mycobacteriaceae sequence relative abundances ............................ 31

TABLE 5 Mycobacteria concentration data compiled .................................... 45

TABLE 6. SSPs considered with respect to NTM ........................................... 61

TABLE S1 qPCR primers, reaction conditions, run information ...................... 69

TABLE S2 Aggregated survey responses and sampling details ......................... 70

TABLE S3 16S rRNA gene amplicon sequences, # returned and # filtered .......... 72

TABLE S4 NTM qPCR and sequencing results by household .......................... 73

TABLE S5 Alpha diversity indices of biofilm sample amplicon sequences .......... 75

TABLE S6 LEfSe results for biofilm sample amplicon sequences .................... 76
LIST OF ILLUSTRATIONS

FIGURE 1. New Jersey physiographic provinces, private wells, and land use .......................... 12
FIGURE 2. NTM marker gene log-relative abundance by home ........................................... 28
FIGURE 3. Heat map, family level, biofilm sample amplicon sequences ............................... 33
FIGURE 4. nMDS plot for biofilm sample amplicon sequences ........................................... 36
FIGURE 5. Overview of systems approach to critical review ................................................. 42
FIGURE 6. Detailed systems concept model ........................................................................ 59
FIGURE S1. Alpha rarefaction curve for biofilm amplicon sequences ............................... 77
FIGURE S2. NTM marker gene log abundance by physiographic region .......................... 78
FIGURE S3. Random forest measures of variable importance ............................................. 79
I. INTRODUCTION

1. Background

Over the past several decades, infections caused by nontuberculous mycobacteria (NTM) have gained increasing attention due to their high cost (Collier et al., 2021), difficult treatment (Daley et al., 2020), and increasing prevalence (Ratnatunga et al., 2020). Given the link between environmental sources of NTM and infections, these trends have highlighted the potential need for engineering controls. NTM, otherwise referred to as “atypical mycobacteria,” environmental mycobacteria, or mycobacteria other than Mycobacterium tuberculosis and Mycobacterium leprae, are ubiquitous in water and soil environments and include many opportunistic pathogens such as those in the Mycobacterium avium complex (MAC) (Falkinham, 2016). Exposures to NTM species by immunocompromised individuals can result in a variety of outcomes including pulmonary disease, as well as extrapulmonary and disseminated diseases (Daley et al., 2020; Falkinham, 1996; Wolinsky, 1979). While the exact number of NTM disease cases in the United States (US) is difficult to ascertain due to lacking reporting requirements (Stout et al., 2016), estimates indicate that cases have increased from approximately 1.8 cases per 100,000 people in the early 1980s (O’Brien et al., 1987) to 16 per 100,000 people in 2014 (Donohue, 2018), increasing at an average rate of 0.44 cases per year. While some of this increase is due to enhanced disease awareness and detection capabilities, researchers believe that an actual increase in cases has also occurred (Kendall and Winthrop, 2013). Cases of NTM infections and disease come with a high price tag. A US Centers for Disease Control (US CDC) study estimated that NTM infections were the costliest primarily water-based infections costing $1.53 billion in direct healthcare costs in the US in 2014 (Collier et al., 2021). For hospitalized patients, NTM infections were the second costliest at $29,600 per episode (Collier et al., 2021). While concerns about NTM disease typically focus on immunocompromised individuals who comprise a large fraction of patients, researchers have recently noted that NTM disease prevalence is increasing in
immunocompetent populations, raising the public health concern around these bacteria (Stout et al., 2016). NTM disease hot spots exist around the US, with a majority of infections occurring in Hawaii, Louisiana, California, Florida, Oklahoma, Pennsylvania, and New York (Adjemian et al., 2012).

There are currently over 190 known NTM species (Daley et al., 2020) that can be divided into two major groups: slow and rapid growing (Stahl and Urbance, 1990). Slow growing species require approximately two to three weeks to culture, while rapid growing species require seven days to culture (Griffith et al., 2007). The most clinically relevant species vary by region across the globe and have changed over time (Hoefsloot et al., 2013). Several slow and rapid growing species are currently of clinical interest. Slow growing pathogenic Mycobacterium species include *M. avium*, *M. intracellular*, *M. kansasii*, *M. marinum*, *M. malmoense*, and *M. xenopi*; and rapid growing species include *M. fortuitum*, *M. abscessus*, and *M. chelonae* (Falkinham, 2013). Another important slow-growing species, *M. gordonae*, is ubiquitous in drinking water environments but is less frequently implicated in disease. These species are generally small (1–3 μm), acid-fast, tubular cells (Glaser et al., 2011) that exhibit varying degrees of resistance to treatments in humans and environmental stresses.

NTM are ubiquitous in diverse environments, including those that may be too harsh for many other microorganisms. A common feature across NTM species that contributes to their success under environmental stresses is the composition of their lipid-rich outer membrane with mycolic acid chains (Brennan and Nikaido, 1995). This outer membrane contributes to cell hydrophobicity, which in turns allows for impermeability and slow growth (Brennan and Nikaido, 1995). Slow growth is also due in part to the low number (one or two) of rRNA copies (Bercovier et al., 1986). Differences in the arrangement of the mycolic acids may contribute to different levels of antibiotic resistance across NTM species (Brennan and Nikaido, 1995). Hydrophobicity
has also been correlated with chlorine resistance, important to NTM’s survival in engineered water systems (Steed and Falkinham, 2006).

NTM’s hydrophobic outer membranes facilitate attachment to other hydrophobic materials (Schulze-Röbbecke et al., 1992). Hydrophobicity enables NTM adsorption to other hydrophobic particles that float in water or that water passes through, such as soils (Bendinger et al., 1993; Brooks, 1983). In one study, this led to a surprisingly low recovery of MAC organisms from soil, where 95% of cells could not be detached from soil particles (Brooks, 1983). However, cells were recovered from clean water that percolated through NTM-contaminated soils, indicating that this exchange may be a transport process that occurs in natural environments, such as during groundwater recharge (Brooks, 1983). In drinking water distribution systems, NTM have been associated with increased turbidity, further supporting the idea that attachment to solids is a transport mechanism in water (Falkinham et al., 2001). In drier environments, adsorption to soils can also result in NTM release with inhalable dust (Bendinger et al., 1993). NTM cells carry a negative surface charge that also contributes to adsorption (Lytle et al., 2004).

Further, NTM cells attach to hydrophobic surfaces, such as pipes, and to other cells, leading to biofilm formation (Schulze-Röbbecke et al., 1992). NTM have frequently been recovered from biofilms in drinking water systems (Falkinham, 2015a), where their concentrations are enriched compared to bulk water (Lehtola et al., 2007). Biofilm cells better resist disinfection, including for a limited period of time after cells are released from the biofilm (Steed and Falkinham, 2006). Part of this increased resistance may be explained by the formation of an extracellular polymeric substance (EPS) matrix that encloses the biofilms (Bardouniotis et al., 2001). Hydraulics can impact NTM in biofilms: Growth of M. avium in biofilms increased with increased flow for one week, but stabilized thereafter irrespective of flow (Torvinen et al., 2007), and steady flow has been associated with growth of several other NTM species (Schulze-Röbbecke et al., 1992). This is attributed to their slow growth and relatively low nutrient demands enabling growth without
direct dependence on flow-related nutrient replenishment (Torvinen et al., 2007). The level of NTM growth in biofilms has unsurprisingly been related to the period of time over which the biofilms develop, and organic substrates, such as polyethylene, were previously found to better support growth of culturable NTM, attributed to their release of nutrients (Schulze-Röbbecke et al., 1992). It has also been suggested that NTM aggregation in biofilms, coupled with the hydrophobic nature of the cell walls, leads to NTM-enriched and easily ejected aerosol droplets (Parker et al., 1983). This is particularly relevant to NTM transport as it relates to human infections, given the potential for inhalation of these aerosols.

Another characteristic that enables NTM’s success in a multitude of environments is the great variability in genotype and phenotype, even within a single species. For example, *M. avium* has opaque and transparent variants with different outer wall properties that the cells are capable of switching back and forth between (Rastogi et al., 1981). The transparent variants have an additional polysaccharide outer wall around the lipid membrane that appears to enhance the disinfectant and antibiotic resistance as well as the virulence of that variant (Rastogi et al., 1981). Some NTM species have been shown to carry plasmids that encode for resistance to heavy metals, such as copper or mercury, and antibiotics (Pal et al., 2015; Ripoll et al., 2009). They can also share this genetic information via plasmid-mediated horizontal gene transfer (Nguyen et al., 2010). With these genes, in addition to membrane impermeability, NTM species such as *M. abscessus* and *M. scrofulaceum* can survive the presence of hydrophilic drugs and heavy metals (Erardi et al., 1987; Jarlier and Nikaido, 1990; Meissner and Falkinham, 1984; Rastogi et al., 1981). Some NTM species also resist ultraviolet (UV) irradiation (David, 1973). In species including *M. avium*, *M. phlei*, and *M. smegmatis*, the production of carotene pigment when exposed to light contributes to UV resistance (Baker, 1938; Tran et al., 2020). These adaptive capabilities of NTM enable their success in several environments and as pathogens.
Many NTM species withstand relatively extreme environments. NTM are considered microaerobic, growing at 6-12% oxygen (Falkinham, 2013; Kirschner et al., 1992), oligotrophic, (Smeulders et al., 1999), resistant to high temperatures (Schulze-Röbbecke and Buchholtz, 1992), and acid tolerant (Bodmer et al., 2000). With respect to oligotrophy, _M. smegmatis_ cultures have been found to survive 650 days of starvation of carbon, nitrogen, and phosphorous, entering into a stationary phase with reduced cell division, stabilized mRNA, and increased resistance to osmotic, acid, and oxidative stresses (Smeulders et al., 1999). Further, NTM species are capable of living on a variety of organic compounds, which also likely enables their success in oligotrophic environments (Falkinham, 2009). Some researchers also suggest that NTM can be carbon-fixing, allowing survival in the absence of nutrients (Zhu et al., 2019). NTM growth is also stimulated by the presence of humic and fulvic acids (Kirschner et al., 1999). With respect to temperature, _M. avium, M. chelonae, M. phlei, M. scrofulaceum, _and M. xenopi_ were all found to resist temperatures as high as 60°C, making them more temperature resistant than _Legionella pneumophila_, the bacteria for which water heater temperature recommendations are designed at 55°C (Schulze-Röbbecke and Buchholtz, 1992). In terms of acid resistance, _M. avium_ has been shown to survive stomach acid and pH as low as 2.2, using its stationary phase to do so (Bodmer et al., 2000). A relationship between lower pH and a greater ability to cultivate NTM is well established (Bodmer et al., 2000; Falkinham, 2009; Kirschner et al., 1992; Robinson, 2019), and an ideal range of 4.5–6.5 has been suggested, based on a study in Hawaii, where the natural environment pH is acidic (Robinson, 2019). These abilities of NTM to survive or thrive under otherwise stressful conditions play an important role as changes to their base levels under climate change stresses are considered.

2. Gaps in understanding and thesis objectives

Much of what is known about NTM in drinking water system biofilms is from the study of public supply systems and premise plumbing (Batté et al., 2003; Haig et al., 2020; Haig et al.,
However, little is known about NTM prevalence in private wells, which nearly 43 million people rely on in the US and that are often used without disinfectants (Johnson et al., 2019). In New Jersey (NJ), where 12% of the population (more than one million people) uses private well water (NJDOH, 2019), microbiological studies have been limited to fecal indicators (Atherholt et al., 2013; Atherholt et al., 2015), although it is known that opportunistic pathogens do not necessarily correlate with these organisms (Wang et al., 2017). In states surrounding NJ (New York and Pennsylvania), NTM disease hot spots have been identified, indicating potential concern for NTM disease in this state given the hydraulic connection (Adjemian et al., 2012).

Also, while much is known about premise plumbing biofilm microbiomes (i.e., homes connected to public water supply) (Donohue et al., 2015; Feazel et al., 2009; Perez-Martinez et al., 2013), little is known about the home biofilm microbiome in homes using private well water (Feazel et al., 2009; Xue et al., 2020). Understanding the biofilm microbiome is important because microorganisms interact with one another, and these interactions carry potential to be exploited for ecological engineering of healthy plumbing microbiomes (Wang et al., 2013a). Thus, the objectives of the first chapter of this thesis were to understand (1) NTM prevalence and abundance in private drinking water wells, and (2) biofilm microbiome composition that may support or limit the success of NTM in this environment.

To meet these objectives, a field study was performed. Biofilm samples were collected from volunteer participant homes of private well owners and from homes of NTM patients as diagnosed by pulmonary doctors in NJ lung clinics. Fecal indicator organisms and water quality parameters were measured, and quantitative polymerase chain reaction (qPCR) was used to quantify NTM gene concentrations in the biofilms. Amplicon sequencing was performed on select samples to analyze the home plumbing microbiome. Survey data was collected from participants about their home plumbing systems that was analyzed for relationships to NTM abundance and to the microbiomes observed. Microbiomes were also analyzed for potential
relationships among organisms that may inform ecological engineering. To the author’s knowledge, this is the first study to investigate NTM in NJ, as well as home plumbing biofilm microbiomes connected to private wells in the state.

As instances of NTM disease have been on the rise (Ratnatunga et al., 2020), and as global climate change exerts numerous effects on natural and engineered systems (IPCC, 2014), there is reason to consider how climate change and associated changes in engineering and societal choices may influence risk of NTM exposure or infection. Several studies have been conducted both at the bench (Brennan and Nikaido, 1995; Lytle et al., 2004; Rastogi et al., 1981; Steed and Falkinham, 2006) and field scale (Delafont et al., 2014; Iivanainen et al., 1993; Kirschner et al., 1992; Schulze-Röbbecke et al., 1992) to understand the influence of environmental conditions on NTM survival and growth. Others have studied associations between climatic and societal factors, and disease risk (Adjemian et al., 2012; Honda et al., 2015; Lipner et al., 2017; Prevots et al., 2014; Thomson et al., 2020). While literature reviews have gathered information pertaining to NTM epidemiology and ecology (Drancourt et al., 2007; Falkinham, 2009; Falkinham, 2013; Honda et al., 2018; Pereira et al., 2020; Thomas and McDonnell, 2007; Vaerewijck et al., 2005), to the author’s knowledge, no comprehensive analysis regarding the potential effects of climate, engineering, and societal change on NTM exposure have been completed. Researchers with the Intergovernmental Panel on Climate Change (IPCC) and beyond suggest that climate and population change together will affect the reliability and quality of water resources (Ashbolt, 2010; IPCC, 2014). Thus, the objectives of the second chapter of this thesis were (1) to understand the interactions occurring at the interface of natural, engineering, and societal systems that may affect NTM exposure, (2) to discuss potential scenarios that might lead to different levels of future risk, and (3) to highlight uncertainties that, if understood, would facilitate more robust analysis of future risk.

To accomplish these objectives, a systematic literature review was performed. Through this process, connections were made between pathogen dynamics, climate factors, engineering
forces, and societal choices, and gaps in understanding were highlighted. Environmental
concentrations of NTM were also collected and assembled for use in future quantitative microbial
risk analyses (QMRA). The purpose of this review is to inform future model-based QMRA, and
to provide insight that may aid water managers in decision-making under changing climates.
II. NONTUBERCULOUS MYCOBACTERIA IN THE BIOFILM MICROBIOME OF PRIVATE WELL AND PREMISE PLUMBING

(Please see Section VII for co-author and pending publication acknowledgement)

1. Background and introduction

Nontuberculous mycobacteria (NTM) are Gram-positive, acid-fast bacteria (Reynolds et al., 2009) ubiquitous in water and soil environments (Falkinham, 2002). Several NTM species are opportunistic pathogens, disproportionately affecting the immunocompromised, resulting in NTM pulmonary disease in those with cystic fibrosis and chronic obstructive pulmonary disease (COPD), as well as elderly populations (Strollo et al., 2015). In recent years, NTM pulmonary disease has also increasingly affected immunocompetent patients, or individuals without classic risk factors (Stout et al., 2016). Cases of NTM pulmonary disease have been on the rise around the world and in the United States (US), where estimates suggest that infection rates nearly doubled from 8.2 per 100,000 persons to 16 per 100,000 persons from 1994 to 2014 (Donohue, 2018; Ratnatunga et al., 2020). NTM pulmonary disease costs in the US increased from $815 million in 2010 (Strollo et al., 2015) to $1.53 billion in 2014 (Collier et al., 2021), with projections to continue increasing. In Hawaii, California, Louisiana, Florida, Oklahoma, and Wisconsin, as well as in Pennsylvania and New York, states directly to the west and north of the location of this study in New Jersey (NJ), Medicare records indicate spatial clusters of NTM pulmonary disease (Adjemian et al., 2012), a trend that has also been observed elsewhere (Lipner et al., 2017). Species responsible for NTM disease vary across the country and globe. In the US, members of the slow-growing *Mycobacterium avium* complex (MAC), *Mycobacterium xenopi*, *Mycobacterium kansasii*, and rapid-growing *Mycobacterium fortuitum* and *Mycobacterium abscessus* are the primary disease causing species (Hoefsloot et al., 2013; Honda et al., 2018). Treatment for NTM pulmonary disease is complicated (Daley et al., 2020), costly (Collier et al., 2021; Strollo et al., 2015), and can involve extended hospital stays in addition to several months of antibiotic therapies (Daley et al., 2020; Dowdell et al., 2019). As NTM infections become
more frequent, and as the population ages, it is of growing importance that environmental exposures to NTM be understood and mitigated.

NTM, as a result of their lipid outer membranes, are hydrophobic (Brennan and Nikaido, 1995), impermeable (Daffé and Draper, 1997; Nikaido et al., 1993), and slow-growing (Daffé and Draper, 1997). They are also capable of hiding within host organisms (Delafont et al., 2014; Dowdell et al., 2019), making them well-suited for growth in biofilms (Falkinham, 2009; Parikh et al., 2019). Studies have shown that NTM are part of the biofilm microbial community in drinking water at treatment plants (Pinto et al., 2012), at points along distribution systems (Haig et al., 2020; September et al., 2004), and at home taps and showerheads (Feazel et al., 2009; Thomson et al., 2013a). Periods of water stagnation allow for NTM proliferation on premise plumbing, which are then released from biofilms on the showerhead or inner faucet into inhalable bioaerosols. Studies have shown that isolates from NTM patients can be clones of isolates from the patient’s home tap, suggesting that infection can come from use of home water and inhalation of bioaerosols (Falkinham, 2013; Falkinham et al., 2008; Thomson et al., 2013a).

A majority of the studies of NTM in the US to date focus on water and biofilms from public water supply distribution systems fed from either surface water (Waak et al., 2019b), groundwater (Klanicova et al., 2013), or both (Falkinham et al., 2001), or on premise plumbing in homes that public systems serve (Donohue et al., 2015; Feazel et al., 2009; Haig et al., 2020; Perez-Martinez et al., 2013). However, 15% of the US population, amounting to 43 million people (Johnson et al., 2019), obtain drinking water through domestic private wells. Domestic private well water systems have seldom been included in the NTM drinking water and household plumbing studies, likely due to reports of the absence of NTM in southeastern groundwaters (Martin et al., 1987). When homes using private wells have been included in the studies, geographical differences in the presence of mycobacteria have been revealed, raising questions about the prevalence of these
microbes in private well water systems elsewhere (Feazel et al., 2009; Gebert et al., 2018; Xue et al., 2020).

The private well microbiome has received considerably less study than the microbiome of public water supplies. Some researchers have investigated groundwater microbiomes (Franca et al., 2015; Smith et al., 2012; Unno et al., 2015) and known contaminated or salt marsh private wells (Pogoda, 2017; Xue et al., 2020). To the author’s knowledge, no studies of private well microbiomes have been performed in NJ where nearly one million people use private wells (NJDOH, 2019). Of further interest when studying the microbial ecology of private wells, which in contrast to public supplies are often not disinfected, is the potential for insight into microbial community members that are or are not associated with microbial agents such as pathogenic NTM. This information could inform ecological engineering solutions to treat undesirable microbial agents in drinking water (Wang et al., 2013a).

To address these gaps in knowledge about private well drinking water systems, we undertook a field study with the following objectives: to quantify NTM in private well household plumbing biofilms as compared to biofilms in NTM patient homes across NJ and to characterize the microbiome in these systems. To achieve these aims, sampling was performed at the homes of participants recruited from the general public and lung clinics in NJ. Biofilm samples were collected from faucets and showerheads and analyzed for NTM and the total microbiome. Water samples and survey data were collected to understand environmental factors potentially influencing the observations. We tested the hypothesis that biofilms from NTM patient homes with public water supply would have higher relative abundances of NTM than biofilms from homes of the general public with private well water. Additionally, we evaluated the hypothesis that one or multiple environmental factors influence NTM prevalence and abundance in biofilms. Finally, we evaluated the hypothesis that biomarker organisms exist that distinguish between the
microbiomes of biofilm samples with and without NTM marker genes, thus serving as potential antagonists or mutualists.

2. Material & methods

2.1. Participant recruitment

Participants were recruited from April 2019 through March 2020. Two populations of individuals were recruited: members of the NJ general public who have private drinking water wells and patients with NTM pulmonary disease as diagnosed by pulmonary doctors according to the criteria set forth by the American Thoracic Society/Infectious Disease Society of America (Griffith et al., 2007), and receiving care at the pulmonary medicine clinics in NJ. Members of the general public were recruited using a variety of outreach efforts including distribution and/or posting of recruitment flyers at local watershed meetings, county extension and local government

**FIG 1** (a) Physiographic regions of NJ, (b) Distribution of private wells in NJ that tested positive for fecal coliform during testing for the NJ Private Well Testing Act between 2002–2014, (c) 2012 generalized land use in NJ (NJDEP, 2020) Disclaimer: These maps were developed using New Jersey Department of Environmental Protection Geographic Information System (NJDEP GIS) digital data, but this secondary product has not been verified by NJDEP and is not state-authorized or endorsed (NJDEP, 2020).
offices, Rutgers campuses, as well as on social media. Patients from participating pulmonary clinics were informed of the study by their doctors and if interested, were followed up with by our team. In total, 19 members of the general public and five NTM patients participated in this study. Note, participant recruitment and sampling efforts were stopped at the onset of the COVID-19 pandemic. Participants completed a survey providing information about their water source, private wells (if applicable), home plumbing and pipe system, and personal habits.

Sampling occurred across the physiographic regions of NJ (Fig. 1). This captured areas of the state that were previously found to have varying degrees of fecal contamination in private well water based on 2002–2014 data represented by the New Jersey Department of Environmental Protection (NJDEP) Geographic Information System (GIS) Private Well Testing Act summary results (Atherholt et al., 2013; NJDEP, 2020).

An IRB protocol was created by the research team and approved by the Rutgers institutional review board. The protocol included a written protocol, phone script, and an informed consent agreement.

2.2. Biofilm collection

Samples collected from each home included two to three biofilm swabs and one water sample (0.5 L). Participants were asked to provide the researchers with access to their most frequently used sinks and showers for sample collection. In total, 70 bulk biofilm samples were collected from participant homes: 56 from members of the general public (19 homes) and 14 from NTM patients (five homes). Biofilm samples were collected for biomolecular analyses of mycobacteria and the total microbial community. Up to three biofilm samples were collected in each home with sterile cotton swabs from the inner faucets of one kitchen, first floor bathroom, or alternative sink, one second floor bathroom sink, and one showerhead. Alternative sinks were either basement or secondary, often less used sinks. To collect biofilm samples from the sinks, water from the cold tap was run for 10 seconds to wet the pipe and aerators were removed using sterile technique
The first six inches of the inside surface of the faucet were swabbed by turning the swab around the faucet for two rotations, and up and down the sides several times to collect as much biofilm as possible. If a sink’s aerator could not be removed and there were no alternative sinks, as was the case in 13 of the 24 homes, the outside of the aerator was swabbed, and the situation was noted. To collect bulk biofilm samples from showerheads, the showerheads were removed, and cool water run before swabbing the inside of the pipe, as described above, then the inside of the showerhead itself was swabbed. [Biofilm thickness was not measured but others reported thicknesses in shower head hoses to range from non-detectable to 0.40 mm (Proctor et al., 2018).] The cotton tips from the swabs were broken off from the wooden stick using ethanol-flame sterilized tweezers. The cotton tips were each stored in a separate sterile 15 mL tube in coolers on ice until returned to the lab for storage at -20°C until DNA extraction. The biofilm field blanks consisted of sterile cotton swabs that were transferred to sterile 15 mL tubes at two participant’s homes (one patient and one general public member) and were otherwise processed the same as the home biofilm samples.

2.3. Water sample collection & quality analyses

Water quality was analyzed in the field and in the laboratory. After collecting the biofilm samples, the faucet and adjacent area were cleaned using 10% bleach, and the water was run at a cool temperature for at least two minutes using only the cold tap (ANRA, 2019). The kitchen sink was used unless the aerator on the faucet was not removable, in which case the water was collected from a first floor bathroom sink. Water temperature was measured using a standard mercury-free glass thermometer placed under a steady stream of cool water. To measure residual chlorine, a low-range (0.2–2.0 mg/L) Hach pocket colorimeter was used with DPD total chlorine reagent packets (Loveland, CO) according to manufacturer’s instructions.

One water sample (0.5 L) per home was collected in a triple-washed and autoclaved polypropylene Nalgene bottle for water quality and fecal indicator organism analyses. Any
residual chlorine in the water sample was quenched with 1% sodium thiosulfate (Allard et al., 2019). Samples were stored in a cooler on ice, as above. The water field blanks consisted of 0.5 L autoclaved deionized water in bottles that were opened at participants’ homes during sample collection and otherwise processed the same as the home water samples. Samples were held no more than three hours in the cooler. Upon returning to the laboratory (~1hr), conductivity and pH were measured using a calibrated Orion Star A329 multimeter (Thermo Scientific, Waltham, MA). US Environmental Protection Agency (USEPA) Method 1604 (USEPA, 2002) was used to evaluate the presence and concentration of total coliforms (TC) and *Escherichia coli* (*E. coli*) in the water samples via plate counts. Two volumes of water (100 mL and 200 mL) per home were analyzed in efforts to obtain countable plates.

2.4. Biomolecular analyses for mycobacteria and microbial community

DNA was extracted from biofilm swabs using a slightly modified phenol-chloroform method PC2 published in 2018 by Haig et al. (Haig et al., 2018) (S1). This method was selected because it was demonstrated to extract DNA more efficiently from a wide range of cells, extracting three times more total DNA and eight times more mycobacterial DNA, on average, than commercial DNA extraction kits. Briefly, cells were lysed with an increased sodium chloride buffer concentration to reduce polysaccharide extraction (Fang et al., 1992). Then, proteins and lipids were density separated from DNA, and DNA extracts were precipitated (Haig et al., 2018). Precipitates were washed with two ethanol wash cycles (Fang et al., 1992).

Quantitative polymerase chain reaction (qPCR) was performed for the 16S rRNA gene and three mycobacterial targets. DNA extracts were diluted in molecular biology grade water (1:10 to 1:100, v/v) to reduce the concentration of PCR inhibitors. The V3 region of the 16S rRNA gene amplified with universal primers was used as a surrogate for the total bacterial population (Muyzer et al., 1993). Mycobacterial targets included were a 16S rRNA genetic marker for *Mycobacterium avium* (*M. avium*) (Wang et al., 2012a), a *Mycobacterium* spp. 16S rRNA genetic
marker (hereafter myco16S) (Radomski et al., 2010), and a mycobacteria functional genetic marker, \( \text{atpE} \), which encodes for the ATP synthase protein subunit C (Radomski et al., 2013). Mycobacteria carry one to two 16S rRNA gene copies (Bercovier et al., 1986), targeted by the myco16S gene, and one of the \( \text{atpE} \) gene (Radomski et al., 2013). For this reason, both genes were measured in all samples, and the results were compared. Excluding \( M. \text{avium} \), a major species responsible for NTM infections, the mycobacteria genetic markers included in this study are specific to all \( \text{Mycobacterium} \) spp., not just NTM. However, with the exception of one study of a wildlife watering hole, \( \text{Mycobacterium} \text{ tuberculosis} \ (M. \text{ tuberculosis}) \) has not been found in water environments outside of a host (Barasona et al., 2017; Kazda et al., 2010). Therefore, general mycobacteria genetic markers likely effectively represent markers for NTM in this environment.

qPCR assays for the V3 variable regions of the 16S rRNA gene and the \( M. \text{avium} \) 16S rRNA gene fragment included 5 µL of SsoFast\textsuperscript{TM} EvaGreen\textsuperscript{R} Supermix (Bio-Rad, Hercules, CA), 0.40 µM forward and reverse primers, 2.4 µL molecular biology grade water, and 1 µL of diluted DNA extract. Assays for myco16S and \( \text{atpE} \) reactions included 5 µL of SsoAdvanced\textsuperscript{TM} Universal Probes Supermix (Bio-Rad, Hercules, CA), 0.65 µM forward and reverse primers, 0.2 µM probe, 1 µL molecular biology grade water, and 1 µL of diluted DNA extract. All qPCR reactions had a total volume of 10 µL per well and were run in triplicate on a Real Time Thermocycler (BioRad CFX96 Touch, Hercules, CA) on 96-well plates. All plates were calibrated with a seven-point calibration curve with concentrations ranging from \( 10^2 \) to \( 10^8 \) gene copies per well. Standards were generated for each primer set with PCR products obtained from either culture streaks of \( M. \text{avium} \) strain ATCC 35717 or from environmental samples and confirmed with Sanger Sequencing (GeneWiz, Piscataway, NJ, available in GenBank under Accession Number MW600431). Briefly, PCR products were cloned with a TOPO TA cloning kit (Invitrogen, Carlsbad, CA) according to the manufacturer’s instructions and quantified according to Pei et al.
A no template control was included in triplicate on each plate. The average $R^2$, run efficiencies, and percent recoveries are provided along with thermocycler conditions and amplicon lengths in Table S1. Amplification specificity was monitored by melt curve analysis for EvaGreen reactions and amplicon lengths were confirmed by gel electrophoresis for select products from all reactions. Gene copies are reported as $\log_{10}(x+1)$ gene copies per swab. The limit of detection (LOD) was determined as the upper limit of the 95% confidence interval established following the procedure from Armbruster and Pry (2008).

For quality assurance and control of the biomolecular methods, blanks and matrix spikes were employed. First, to evaluate contamination throughout the extraction process, a total of four DNA extraction process blanks were performed. These blanks consisted of sterile cotton swabs dipped in molecular biology grade water. Extracts from these swabs were then tested for contamination by qPCR for the presence of target genes and quantities of 16S rRNA gene copies. Matrix spikes were performed with isolates of *M. avium* as the matrix spike. Briefly, *M. avium* colonies were scraped and boiled in 10 μL water to elute the DNA. Eluted DNA was then diluted in water 1:50, and 10 μL of the diluted DNA were spiked into two test swabs, one with molecular biology grade water, and one with biofilm from a laboratory faucet. A replicate swab from the laboratory faucet served as a biofilm control that was not spiked with *M. avium*. qPCR on the resulting extracts targeting *M. avium* showed that similar amounts of *M. avium* were extracted from both the spiked sterile water sample and the spiked sink biofilm sample (4.80 log copies and 4.44 log copies, respectively, relative percent difference on log scale of 7.8%), while the negative control sample had no detectable *M. avium*.

To investigate prokaryotic diversity of the biofilm samples, Illumina MiSeq amplicon sequencing (300bp, paired end) was performed on a subset of samples targeting the V3-V4 region of the 16S rRNA gene at a commercial laboratory (MR DNA, Shallowater, TX). A total of 61 samples comprised of 11 duplicate sets and 13 triplicate sets (i.e., samples from different
locations in the same home) were sent for sequencing analysis. The subsample set was chosen to provide a range of pipe materials, physiography, and mycobacteria detections (Table S2).

Sequences were processed in Qiime2-2019.10 on the Rutgers Office of Advanced Research Computing (OARC) Amarel computer cluster following a protocol established from the Qiime2 “Atacama soil microbiome” tutorial (Bolyen et al., 2019) and a pipeline published in 2019 (Payne et al., 2019). Briefly, sequences were demultiplexed, trimmed, and processed through the DADA2 package which filtered, denoised, and removed chimeras (Callahan et al., 2016).

Taxonomy was assigned using Silva 132 reference taxonomy alignment with the 99% taxonomy set and classified with the Naïve Bayes classifier clustered at 97% similarity. A rarefaction subsampling depth of 22,000 sequences per sample was selected, which included 29 samples (see Table S3 for details). Notably, all of the showerhead samples sequenced returned less than 100 raw sequences each and thus were not included in downstream analyses. These showerhead samples comprised 72% of the 32 samples excluded from the rarefied analyses. With these samples excluded, the sequences analyzed were comprised of nine sets of duplicates (18 samples, two per home) and 11 samples from individual homes. As seen in the alpha rarefaction curve (Fig. S1), the rarefaction subsampling depth of N = 22,000 was sufficiently deep to capture the diversity.

2.5. Statistical analyses

All statistical analyses were performed in R version 4.0.0 (http://www.rproject.org). Data normality was assessed using the Shapiro-Wilk test. Differences in water quality across physiographic regions and water supplies were evaluated using pairwise nonparametric permutational multivariate analysis of variance (pairwise PERMANOVA) tests with a Bonferroni

\[ P \text{-adjustment for multiple comparisons via PairwiseAdonis version 0.3 (Martinez Arbizu, 2020).} \]

Differences in marker gene quantities between private well and public supply samples were evaluated with a nonparametric Kruskal-Wallis test followed by a Wilcoxon rank sum test. Due to a high percentage of samples (e.g., 75% for \text{atpE}) having quantities of marker genes below the
LOD or limit of quantitation (LOQ), application of methods for left-censored data were necessary. For the quantitative analysis, values below the LOD were substituted with LOD/√2 and LOQ were substituted with LOQ/√2 (Ganser and Hewett, 2010). To estimate the average abundances for each physiographic region and water supply, nonparametric Kaplan-Meier (K-M) estimation was performed with both myco16S and atpE log-transformed abundances (Antweiler and Taylor, 2008; Helsel, 2012). The K-M estimation for left-censored data was performed using the NADA R package version 1.6-1.1 (Lee, 2020). To evaluate the importance of the variables studied on the total 16S rRNA gene copies and relative concentrations of myco16S genetic markers, Random Forest (RF) models were developed using the randomForest R Package version 4.6.14. (Liaw and Wiener, 2002). This ensemble machine learning algorithm is driven by decision trees, is nonparametric, and has many strengths including the ability to analyze both categorical and continuous variables, manage missing data, and analyze small datasets with many predictors (Ali et al., 2012; Biau and Scornet, 2016; Breiman, 2001; 2002; Luan et al., 2020). RF uses bootstrapping to grow unpruned decision trees to train the model, and ‘out of bag’ data (approximately 1/3 of all data) are used to calculate error, eliminating the need for a separate training and test set (Breiman, 2001; Liaw and Wiener, 2002). Following the guidance of Mendez and Lohr (2011), the regression models were calculated and adjusted by removing variables that produced a negative percent increase in mean square error (%IncMSE) until all %IncMSE values were positive (Mendez and Lohr, 2011). Because RF models may underestimate variable importance with co-correlated variables (Biau and Scornet, 2016), one variable from each pair of predictors that had a Spearman’s rho coefficient > 0.50 was used per model. To better understand the variation in factors included in the optimized RF models, pairwise PERMANOVA analyses were subsequently performed on these variables using PairwiseAdonis version 0.3 (Martinez Arbizu, 2020).
Alpha and beta diversity of amplicon sequences were analyzed at the family level subsampled at N = 22,000 sequences per sample. Alpha diversity indices calculated included species richness, Shannon’s $H'$, and Shannon’s evenness. Indices were compared between private well and public supply samples using a Wilcoxon rank sum test. The relative importance of potential explanatory factors (i.e., survey data and water quality) to the observed variation of alpha indices was evaluated using RF models using the same methodology explained above. For beta diversity, a Bray-Curtis Dissimilarity matrix was calculated using the “vegan” R package version 2-5.6 (Oksanen et al., 2019) with log$_{10}$(x+1) transformed sequence data. Differences in the community structure as a function of potential explanatory variables were evaluated using the same pairwise PERMANOVA tests described for water quality comparisons above using PairwiseAdonis version 0.3 (Martinez Arbizu, 2020). Nonmetric multidimensional scaling (nMDS) plots were generated using the same Bray-Curtis Dissimilarity matrix in the “ggplot2” package in R to view potential clusters (Wickham, 2016).

Linear discriminant analysis (LDA) effect size (LEfSe) tests were conducted using the method published by Segata (2011) with default parameters to identify potential prokaryotic biomarkers that distinguish between the microbiomes of samples from different categories. While LEfSe is traditionally used to identify biomarkers of descriptive or treatment type groupings, here the technique was used to identify potential biomarkers that could inform antagonist or mutualist relationships with NTM by grouping samples by those with and without detection of myco16S (Puzon et al., 2017). Additionally, the method was used in its more typical manner to assess for potential biomarkers between samples from homes with private well water compared to samples from homes with public water supply. These categories were chosen because previous studies have found differences in microbial communities to be most related to water treatment or water source (Ji et al., 2015; Wang et al., 2014). Amplicon sequences are available in the NCBI SRA database (https://www.ncbi.nlm.nih.gov) under Accession Numbers SRR13214198.
3. Results & discussion

3.1. Water quality

**TABLE 1** Water quality results for private well water samples by physiography and public water supply samples. \( N = \) the number of water samples taken per region/water supply. One water sample was taken per home. Average values for water quality parameters are reported ± standard deviation.

<table>
<thead>
<tr>
<th>Region / Source</th>
<th>Water samples (N)</th>
<th>Number of Patients</th>
<th>pH</th>
<th>Conductivity ((\mu)S/cm)</th>
<th>(\text{Cl}_2) (mg/L)</th>
<th>Temperature (°C)</th>
<th>Coliform # (+) homes</th>
<th>E. coli # (+) homes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal Plain</td>
<td>5</td>
<td>0</td>
<td>7.94 ± 0.29</td>
<td>247 ± 87</td>
<td>0.02 ± 0.03</td>
<td>16.3 ± 1.8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Piedmont</td>
<td>11</td>
<td>1</td>
<td>7.18 ± 0.43</td>
<td>656 ± 389</td>
<td>0.04 ± 0.03</td>
<td>20.6 ± 7.4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Highlands</td>
<td>3</td>
<td>0</td>
<td>7.18 ± 0.67</td>
<td>487 ± 164</td>
<td>0.02 ± 0.01</td>
<td>19.4 ± 3.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Valley &amp; Ridge</td>
<td>1</td>
<td>0</td>
<td>7.27</td>
<td>815</td>
<td>0.02</td>
<td>17.0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Public supply</td>
<td>4</td>
<td>4</td>
<td>7.60 ± 0.23</td>
<td>573 ± 74</td>
<td>1.02 ± 0.19</td>
<td>15.8 ± 5.0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Water quality and survey data were collected to evaluate potential relationships with the microbial communities studied. Private well water pH and conductivity were found to vary across NJ’s physiographic regions (Table 1). In the Coastal Plain, pH was significantly higher \((P = 0.01, \text{ pairwise PERMANOVA})\) compared to the Piedmont and Highlands regions. pH in the public water supply was higher than in the Piedmont region \((P = 0.04, \text{ pairwise PERMANOVA})\) and similar to pH in well water from other regions. Note, given that only one home in the Valley and Ridge region was sampled, robust comparisons for that region are not possible. The pHs observed fell within or near to the recommended pH range (6.5–8.5) for public water supplies (NJ Administrative Code § 7:10-7.2.) Conductivity in the Coastal Plain was significantly lower than in the Piedmont and Highlands regions \((P = 0.01, \text{ pairwise PERMANOVA})\). The NTM patient
home public supply water had similar conductivity to the Piedmont and Highlands regions \( (P > 0.77, \text{ pairwise PERMANOVA}) \) and higher conductivity than the Coastal Plain \( (P = 0.01, \text{ pairwise PERMANOVA}) \). These differences for the well water are not surprising given that groundwater chemistry is a function of the minerals, soil types and bedrocks it flows through in the different physiographic regions (Serfes, 2004). As expected, higher residual chlorine was observed in the public supply than in private well water from all regions \( (all \ P = 0.01, \text{ pairwise PERMANOVA}) \), which were similar to one another \( (P > 0.55, \text{ pairwise PERMANOVA}) \). Residual chlorine measurements in public supply samples were, on average, 35 times higher than those of private well water samples, but all measurements were below the Maximum Residual Disinfectant Level Goal of 4 mg/L as Cl\(_2\) (USEPA, 2020).

Cultivable fecal indicators were rarely observed in the well water samples. Total coliforms \( (\text{TC}, 1–10 \text{ CFU/100mL}) \) were observed in three well water samples \( (N = 3/20, 15\%) \) and \( \text{E. coli} \) colonies \( (\text{EC}, \text{between} \ 2–7 \text{ CFU/100 mL}) \) were observed in two of these samples \( (N = 2/20, 10\%) \). Participants were informed of these results and directed to NJDEP guidelines suggesting re-testing of well water with fecal indicators and mitigation based upon those results. Fecal indicators were not observed in any samples from the public supply \( (N = 4) \). Data collected for the NJ Private Well Testing Act indicated between 0–20\% of private wells tested in 2-mile by 2-mile testing grids were positive for fecal coliforms between 2002 to 2014 (NJDEP, 2020). In those data, fecal contamination was least frequent in the Coastal Plain region (Atherholt et al., 2013; NJDEP, 2020), consistent with our results where fecal indicators were not observed in wells from this region. Similarly, results of a Virginia-based study indicated that prevalence of TC and EC was lowest in private well water samples from the Coastal Plain and higher in the Blue Ridge-Piedmont and Valley and Ridge regions (Pieper et al., 2016).
3.2. Survey results

Participants answered questions about their home plumbing systems, including treatment systems and pipe materials (Table S2). Most homes with private wells (N = 17/20, 85%) used a treatment system, with many homes combining multiple treatment systems (information on treatment system age and maintenance was not gathered). Several homes used a water softener (N = 4), an in-line filter (N = 3), or both (N = 7). Three homes with private wells had UV disinfection systems and one used chlorine to disinfect. Of the four patient homes connected to public water supplies, two received chlorinated surface water, one chloraminated surface water, and one chloraminated groundwater, each from a different treatment plant. All participants but one reported their home pipe materials: 13 had copper pipes (54%), seven had mixed copper and PVC (29%), two had only PVC (9%), and one had mixed PVC and PEX (4%).

Information specific to private wells was also collected including well age and construction method. The private wells sampled had a wide range of ages, with the average being 43 ± 20 years. Most participants did not know the method of their well’s construction nor its depth (N = 14, 13 respectively). For those who did know, their wells were drilled (N = 5) or bored (N = 1), and the average well depth was 188 ± 100 ft. A majority of the homes with private wells also had septic tanks (N = 16/20).

The ability of the homeowners to answer questions related to knowledge of basic well maintenance and conditions varied. With regards to water quality testing, 12 participants had testing in the last five years, five had testing before that, and two participants had never had it tested. Regarding maintenance, 13 participants had some form of maintenance performed in the last 10 years, one had maintenance performed before that, and six participants did not know when the last maintenance had occurred. For factors known to be potentially related to contamination, most participants reported that their wellheads were visible aboveground (N = 16/20), and four reported that the soil around that visible wellhead was raised. One participant reported knowledge
about a crack in the casing of their well, while the remaining participants either reported no
cracks or that they did not know. As a previous study of private well water noted, a lack of
knowledge by private well homeowners about their well characteristics is not uncommon (Pieper
et al., 2016).

3.3. qPCR for mycobacteria and total bacterial community

qPCR was performed on all biofilm samples to measure NTM marker genes per swab and
their relative abundances (i.e., gene copies/16S rRNA gene copies). Of the 70 biofilm samples
analyzed, 42 were positive for a myco16S (60%), of which 21 were above the limit of
quantitation (LOQ, 29%). Nineteen of the 70 samples were positive for the atpE gene (27%), of
which 16 were above the LOQ (23%). Concentrations of the myco16S gene moderately
correlated with those of the atpE gene (Spearman’s rho = 0.42, $P = 2.6 \times 10^{-4}$), and the 16S
rRNA gene (Spearman’s rho = 0.59, $P = 7.2 \times 10^{-8}$). For the private well biofilm samples (N =
59), myco16S was detected in 58% and atpE in 30% (Table 2). For the public supply biofilm
samples (N = 11), 64% were positive for myco16S and 9% for the atpE gene. Half of NTM

<table>
<thead>
<tr>
<th>Physiographic region</th>
<th>Number of biofilm samples collected</th>
<th>% positive (^a) for atpE</th>
<th>% positive (^a) for myco16S</th>
<th>average log gene copies / swab (atpE) (95% \text{ CI})(^b)</th>
<th>average log gene copies / swab (myco16S) (95% \text{ CI})(^b)</th>
<th>Average log-relative abundance (atpE) (^c)</th>
<th>Average log-relative abundance (myco16S) (^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal Plain</td>
<td>15</td>
<td>13</td>
<td>40</td>
<td>-</td>
<td>5.21 [4.93, 5.49]</td>
<td>-1.32</td>
<td>-1.55 ± 0.19</td>
</tr>
<tr>
<td>Piedmont (^d)</td>
<td>33</td>
<td>39</td>
<td>73</td>
<td>5.04 [4.80, 5.28]</td>
<td>5.04 [4.76, 5.32]</td>
<td>-1.93 ± 0.58</td>
<td>-2.1 ± 0.90</td>
</tr>
<tr>
<td>Highlands</td>
<td>8</td>
<td>38</td>
<td>50</td>
<td>4.91 [4.48, 5.33]</td>
<td>-</td>
<td>-2.21 ± 0.11</td>
<td>-3.15 ± 1.4</td>
</tr>
<tr>
<td>Valley &amp; Ridge</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>&lt; LOD</td>
<td>&lt; LOD</td>
<td>&lt; LOD</td>
<td>&lt; LOD</td>
</tr>
</tbody>
</table>

TABLE 2 NTM marker gene presence/absence, Kaplan-Meier (K-M) quantitation estimates, and
log-relative abundances for samples with quantities above the limit of quantitation for private
well samples across physiographic regions in NJ. K-M average estimates are reported as average
[lower limit, upper limit] of the 95% confidence interval (CI). Average relative abundance values
are reported ± standard deviation.
a Positive defined as concentration > LOD

b Kaplan-Meier estimated quantities are included for groups with ≥ 2 quantifiable samples, or the minimum required to produce an estimated average with a CI

c Only quantifiable samples are reported; if listed without ‘±’, one sample was quantifiable

d Patient private well samples (N = 3) are included under the Piedmont region

pulmonary patient biofilm swabs (N = 14, majority public supply), were positive for myco16S, and one was positive for atpE (Table 3). These results are similar to a recent study investigating Louisiana private well water using qPCR, which found mycobacteria in 68% of samples (Xue et al., 2020). *M. avium* was not detected in any samples from our study. Matrix spikes performed with *M. avium* resulted in 98.3 +/- 3.1% recovery, suggesting that inhibition was likely not the cause of this non-detection. One study of NTM in unchlorinated drinking water in the Netherlands and another study in the US of drinking water and biofilms, disinfected and not, similarly did not observe *M. avium* when using qPCR targeting *M. avium* (van der Wielen and van der Kooij, 2013; Waak et al., 2019b).

**TABLE 3** NTM marker gene presence/absence, Kaplan-Meier (K-M) quantitation estimates, and log-relative abundances for samples with quantities above the limit of quantitation by water supply for NTM pulmonary patients in NJ. K-M average estimates are reported as average [lower limit, upper limit] of the 95% confidence interval (CI). Average relative abundance values are reported ± standard deviation.

<table>
<thead>
<tr>
<th>Water supply</th>
<th>Number of biofilm samples collected</th>
<th>% positive a for atpE</th>
<th>% positive a for myco16S</th>
<th>average log gene copies / swab atpE b [95% CI]</th>
<th>average log gene copies / swab myco16S b [95% CI]</th>
<th>Average log-relative abundance atpE c</th>
<th>Average log-relative abundance myco16S c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private well d</td>
<td>3</td>
<td>0</td>
<td>67</td>
<td>-</td>
<td>-</td>
<td>&lt; LOD</td>
<td>&lt; LOQ</td>
</tr>
<tr>
<td>Public supply-</td>
<td>9</td>
<td>11</td>
<td>78</td>
<td>5.06 [4.94, 5.19]</td>
<td>-2.60</td>
<td>-2.95 ± 0.75</td>
<td></td>
</tr>
<tr>
<td>surface water</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public supply-</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>&lt; LOD</td>
<td>&lt; LOD</td>
</tr>
<tr>
<td>groundwater</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
a Positive defined as concentration > LOD

b Kaplan-Meier estimated quantities are included for groups with ≥ 2 quantifiable samples, or the minimum required to produce an estimated average with a CI

c Only quantifiable samples are reported; if listed without “±”, one sample was quantifiable

d Private well patient samples are also included in Table 3 in the Piedmont region.

Likewise, another study investigating mycobacteria in showerhead microbiomes in the US did not detect *M. avium* by qPCR in four homes with private wells nor mycobacteria (Feazel et al., 2009), the latter in contrast to our findings. While an exact connection between NTM patient home samples and NTM disease was beyond the scope of this study, the lack of *M. avium* in this study’s samples indicates that *M. avium* in home biofilms may not be the source of their infections. This could imply that other mycobacteria species, such as *Mycobacterium abscessus* or *Mycobacterium kansasii* (Thomson et al., 2020), are responsible for infection and subsequent disease, or that disease was obtained from a different transmission route, such as by inhalation of contaminated household dust (De Groote and Huit, 2006).

For samples with *atpE* or *myco16S* above the LOQ, the average log-relative abundance (relative to 16S rRNA gene copies) was calculated. By this method, only *myco16S* log-relative abundances could be compared across physiography (Coastal Plain and Piedmont, Table 2) and water supply. The average log-relative abundance was highest in the Coastal Plain and lowest in the public supply biofilm samples, corresponding to the majority of patient home samples (N = 11/14), though differences were not statistically significant (*P* = 0.11, Kruskal-Wallis). The Piedmont region had the most biofilm samples with marker gene abundances above the limit of quantitation (Fig. S2). For both NTM marker genes, the abundance ranged from below detection to approximately seven log gene copies per swab. A recent study of NTM in municipal, surface water-sourced drinking water and associated biofilms found similar quantities of the *atpE* gene
in biofilm samples, ranging from below quantitation to $10^6$ copies per cm$^2$ (Waak et al., 2019b), which translates to approximately half the area of one swab in this study. Although the marker gene concentrations are similar between this study and that by Waak et al. (2019b), differences could be explained by differences in biofilm sampling techniques for which no common best practice exists (Wang et al., 2017). Unlike the findings of previous studies (Feazel et al., 2009; Gebert et al., 2018), our study does not show a higher abundance of mycobacteria in public supply biofilms than in those from private wells ($P > 0.16$, Kruskal-Wallis), potentially as a result of the relatively few public supply samples included here.

Of the 24 homes sampled, half had at least one sample that was positive for $atpE$ and all had at least one sample that was positive for myco16S. For patient homes specifically, only one home was positive for $atpE$, and four out of five were positive for myco16S. These results indicate that mycobacteria are present in most homes, and that the myco16S gene was more sensitive than the $atpE$ gene for mycobacteria detection. This could be because there are more (one or two) 16S rRNA gene copies per mycobacterial cell than there are $atpE$ gene copies (one) (Bercovier et al., 1986). This could also be in part because the $atpE$ gene had a higher LOD. Biofilm samples from the same home did not necessarily have similar relative abundances of the target genes. For example, the variance of the log-relative abundance for $atpE$ within one home ranged from 0.03 to 5.0 (average $1.7 \pm 1.7$-log) and for myco16S ranged from $2.4 \times 10^{-3}$ to 4.1 (average $1.7 \pm 1.4$-log). The location within the home with the highest relative abundance was not consistent: showerheads had the highest relative abundances of $atpE$ in 7/24 homes and myco16S in 10/24 homes. Among sinks, second floor bathroom sinks more frequently had the highest relative abundance of $atpE$, while the same was true of first floor bathroom sinks for myco16S (Fig. 2). A previous culture-based study suggested that different mycobacterial species tend to colonize biofilms in some home locations, such as kitchen sinks and showerheads, more frequently than
others (Honda et al., 2016). A species-level analysis of our data was not possible with the techniques applied here.

**FIG 2.** NTM marker gene relative abundances for (a) *atpE* and (b) *myco16S* by home, marked with level of quantitation for each sample collected in a home of the general public (WG#) or an NTM patient (WN# or PN#, for private wells or public supply). Point color corresponds to whether samples were quantified or substituted with LOD/$\sqrt{2}$ or LOQ/$\sqrt{2}$ for NTM marker gene abundances below the LOD ($<$ LOD) or below LOQ and above the LOD ($<$ LOQ). Shapes correspond to location within the home the samples were collected with bathroom 1 a first floor bathroom sink, and bathroom 2 a second floor bathroom sink. Alternative sinks were basement or secondary sinks. Box color corresponds to whether the home was that of an NTM pulmonary patient, or a member of the general public.
No significant differences were observed between the private and public supply biofilm 16S rRNA gene concentrations \( (P = 0.20, \text{ Wilcoxon}) \). The quantities observed here are similar to or slightly higher than those reported in another study of surface water-sourced drinking water biofilm samples in water distribution systems (Waak et al., 2019b). Three biofilm samples from two homes had quantities of 16S rRNA gene copies below the LOD. NTM marker genes were also not identified in these samples, indicating that low biomass may have caused the non-detection in these samples.

### 3.3.1. Random forest regression

Random Forest (RF) regressions were performed to understand the importance of potential explanatory variables to the relative abundances of NTM marker gene myco16S and 16S rRNA gene copies in private well biofilm samples. The RF regression calculated for the log-relative abundance of myco16S genes was optimized to explain 21% of the variation observed with three predictor variables: physiographic region, within-home location, and the presence of raised soil around the wellhead (Fig. S3a). Others have found geography to relate to mycobacteria abundance (Gebert et al., 2018) and soil type to relate to NTM disease risk (Lipner et al., 2017). The high relative importance of within home location may indicate that the environment of the faucet, such as humidity which one might expect to be higher in bathrooms than in kitchens, may influence the relative abundance of mycobacteria. Previous work that has found correlations between humidity and culturable mycobacteria abundance supports this idea (Kirschner et al., 1992). The final variable in the model, the presence of soil around the wellhead, appears to be driven by a higher relative abundance of myco16S genes in samples from wells with soil around the wellhead as compared to samples from participants who did not know the answer to this question. Therefore, this variable, although useful for the model’s predictive capabilities, cannot be meaningfully interpreted. For \textit{atp}E, RF models were not developed, as more than 75% of private well biofilm samples had concentrations below the LOQ, to prevent potential bias from
the abundance of substituted/ censored data (Antweiler and Taylor, 2008). Likewise, due to the low number of public supply biofilm samples collected (N = 11), and the even lower number that had concentrations of the target genes above the LOQ, RF models were not developed for these systems. Previous work regarding the factors that influence mycobacterial abundances in public drinking water biofilms have shown that disinfectant type and residual concentration (Gomez-Smith et al., 2015; Waak et al., 2019a; Waak et al., 2019b; Wang et al., 2012a), flow velocity (Douterelo et al., 2017), water age and pipe material, as well as interactions between these factors (Wang et al., 2014), can influence observed quantities.

Using 16S rRNA gene copies collected per swab as a surrogate for total bacterial population, 27% of the variation in total bacterial abundance in private well biofilms was explained by two predictor variables: pipe material and physiography (Fig. S3b). Interestingly, biofilms from homes that used copper pipes had significantly higher concentrations of 16S rRNA gene copies than biofilms from homes using PVC ($P = 0.04$, pairwise PERMANOVA). This could be due to higher surface area in the copper pipe systems as a result of corrosion, which PVC pipes are not subject to (Wang et al., 2014). Previous work also found that pipe materials significantly influenced the amount of 16S rRNA gene copies in drinking water biofilms (Wang et al., 2014). Physiography may influence the abundance of 16S rRNA gene copies for similar reasons as for mycobacteria. The geophysical properties of the soil that water flows through may facilitate different levels of filtration and adsorption, therefore influencing the concentrations of 16S rRNA genes present in the water and subsequently in the biofilms (Lipner et al., 2017).

3.4. Biofilm microbiome

*Mycobacteriaceae*, the microbial family containing only the genus *Mycobacterium*, were observed in the amplicon sequencing results for 29 biofilm samples from 20 homes in concentrations ranging 0.14–50.6% of the microbiome (log-relative abundance: -2.85–0.30) (Table 4). When sequencing was performed on two biofilm samples from the same home, the
relative percent difference of log-relative *Mycobacteriaceae* abundances from different locations [100x |Sample 1 – Sample 2|/(Average of Sample 1 and Sample 2)] was between 1.81–141%, similar to that of qPCR, which for the same samples was 2.23–104%. Similar to the results of qPCR, the location with the highest log-relative abundance of *Mycobacteriaceae* observations within a home was inconsistent: upstairs bathroom sink was greatest for 5/9 homes, downstairs bathroom sink for 2/9 homes, and kitchen sink for the other 2/9 homes. As with qPCR, log-relative abundance of *Mycobacteriaceae* was highest in the Coastal Plain region, though the difference was not statistically significant (\( P = 0.80 \), Kruskal-Wallis).

The log-relative abundance of myco16S was moderately correlated with log\(_{10}\) transformed relative abundance of *Mycobacteriaceae* observations in sequenced samples (Spearman’s rho = 0.60, \( P = 8.1 \times 10^{-4} \), Table S4). There are over 190 species of *Mycobacterium* (Parte, 2014), and the myco16S primers used in this study have only been tested with 30 of the species and found to have 77% sensitivity and 100% specificity (Radomski et al., 2010), while the *atpE* gene has been tested for sensitivity and specificity (100% for both) with 31 of these species (Radomski et al., 2013). Thus, species not targeted by the primers but identified in amplicon sequences may account for some of the differences observed between these results.

**TABLE 4** *Mycobacteriaceae* sequences for 29 samples by physiography and water supply.

Values reported as average ± standard deviation. *Mycobacteriaceae* sequences were observed in every sample for which sequences were analyzed.

<table>
<thead>
<tr>
<th>Physiographic region</th>
<th>Number of biofilm samples analyzed</th>
<th>Log-relative abundance</th>
<th>% of total observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal Plain</td>
<td>5</td>
<td>-1.71 ± 1.26</td>
<td>18.7 ± 25.5</td>
</tr>
<tr>
<td>Piedmont a</td>
<td>10</td>
<td>-2.08 ± 0.96</td>
<td>8.01 ± 16.1</td>
</tr>
<tr>
<td>Highlands</td>
<td>4</td>
<td>-2.21 ± 0.67</td>
<td>1.77 ± 2.93</td>
</tr>
<tr>
<td>Valley &amp; Ridge</td>
<td>2</td>
<td>-2.60 ± 0.07</td>
<td>0.255 ± 3.86×10(^{-2})</td>
</tr>
<tr>
<td>Public Supply-SW b</td>
<td>6</td>
<td>-2.16 ± 0.73</td>
<td>3.70 ± 8.72</td>
</tr>
<tr>
<td>Public Supply-GW b</td>
<td>2</td>
<td>-2.56 ± 0.13</td>
<td>0.283 ± 8.51×10(^{-2})</td>
</tr>
</tbody>
</table>

\(^a\) Includes two NTM pulmonary patient samples

\(^b\) NTM pulmonary patient samples; SW= surface water sourced; GW= groundwater sourced
In terms of the microbiome composition, *Mycobacteriaceae* was one of the top three most abundant families in six private well biofilm samples and one public supply biofilm sample, comprising an average of 8.6 ± 17% of the private well and 3.3 ± 6.8% of public supply biofilm sample observations (Fig. 3). Although species level short-read OTU classifications should be interpreted cautiously, 49 ± 25% of observations were further classified (using Silva 132 reference taxonomy) as belonging to the species *Mycobacterium gordonae* (*M. gordonae*), a ubiquitous saprophytic species that has exhibited human pathogenicity with increasing frequency in recent years (Mazumder et al., 2010; Utsugi et al., 2015; Zlojtro et al., 2015). Other studies using a combination of culture and biomolecular techniques have similarly found *M. gordonae* to be the most prevalent NTM species in groundwater sourced drinking water and biofilm samples from community water supplies (Covert et al., 1999; Perez-Martinez et al., 2013; September et al., 2004). Additionally, a few observations each were further classified as *Mycobacterium abscessus* (*M. abscessus*) and *Mycobacterium xenopi* (*M. xenopi*), both of which are known human pathogens (Wallace et al., 1998). Previous research of a surface water-sourced, chloraminated distribution system similarly found *M. abscessus* and *M. xenopi* as a small percentage of observations in drinking water biofilm samples (Gomez-Smith et al., 2015). Again, these species level observations are reported to provide suggestions for potential targets in future studies involving private well biofilms and require confirmation with other methods.
FIG 3 Heat map of family level OTUs > 1% of all observations and biomarkers (N = 22000 sequences per sample) collected from sinks in the homes of the general public (WG#) or an NTM patient (WN# and PN#, for private wells and public supply). Samples are arranged according to hierarchical clustering with a Bray-Curtis Dissimilarity matrix and complete clustering method, as indicated by the top dendrogram. The row dendrogram (left) indicates clustering by row means (i.e., similarity of mean observations for each OTU). OTUs are listed as family [class]. Sample names coded as W = private well, G = general public, P = public supply, N = patient; b1 = first floor bathroom sink, b2 = second floor bathroom sink, k = kitchen sink, a = alternative sink, remaining 1–2 digits = system ID (1–20 for private wells; 1–4 for public supply). Note: * potential mycobacteria antagonists, **biomarkers for private well samples, *** biomarkers for public supply samples.

Of the four samples with the highest relative abundances of *Mycobacteriaceae* sequences (27–51%), three were from homes with copper plumbing systems, and one used mixed copper and PVC. Previous work has suggested that, unlike many other microbes, mycobacteria carry
plasmids that encode resistance to copper, potentially explaining the relatively higher abundance of mycobacteria in these samples with copper plumbing systems (Erardi et al., 1987). The samples also reflected some of the highest myco16S relative abundances in the qPCR data.

3.4.1. Total microbial community analyses via amplicon sequencing

Amplicon sequencing was performed also to characterize the microbial communities more broadly in biofilm samples from homes served by private wells and homes of NTM patients served by public water supply. Of the 70 biofilm samples collected, 29 had sufficient sequences to be fully analyzed (Table S3).

At the family level, 21 bacterial families comprised more than 1% of all observations (Fig. 3). For 13 private well and five public supply biofilm samples, the Sphingomonadaceae family ranked as one of the top three most abundant families present per sample, comprising an average of 14 ± 13% and 17 ± 14% of observations for private well and public supply samples, respectively. Members of the alphaproteobacterial Sphingomonadaceae family have been frequently found in drinking water and associated biofilms and were seen to exhibit resistance to disinfection and antibiotics (Idi et al., 2010; Narciso-da-Rocha et al., 2014; Vaz-Moreira et al., 2011). Some members of the Sphingomonadaceae family, such as Sphingomonas paucimobilis, are potential opportunistic pathogens and have been implicated in nosocomial disease (Narciso-da-Rocha et al., 2014; Özdemir et al., 2011; Ryan and Adley, 2010).

Measures of diversity within each sample (alpha) and across all samples (beta) were calculated and analyzed to understand potential explanatory factors. Alpha diversity indices calculated were species richness, Shannon’s $H'$, and Shannon’s evenness. Species richness was significantly higher in private well biofilms than in biofilms collected from homes on public supply ($P = 2.6 \times 10^{-6}$, Wilcoxon, Table S5). Shannon’s $H'$ and evenness were similar for private well systems and public supply, alike ($P > 0.24$, Kruskal-Wallis). The influence of potential explanatory variables on the values of alpha indices was assessed using RF regression models.
For Shannon’s $H'$ (37% variation explained) and evenness (30% variation explained), the RF model found that the location within the home was the most important predictor variable (Fig. S3). This appears to be driven by the lower $H'$ ($P = 0.024$, pairwise PERMANOVA) and evenness ($P = 0.006$, pairwise PERMANOVA) in first floor bathrooms compared to second floor bathrooms, potentially indicative of different levels of use and environmental conditions between the lesser used first floor bathrooms and second floor bathrooms that frequently contained the home’s primary shower. The RF model for species richness explained a lesser 20% of the variation with the visibility of a participant’s wellhead aboveground as the most important variable. Interestingly, richness was nearly double in samples whose wellheads were not visible aboveground versus those whose wellheads were. For beta diversity analyses, 2D nMDS plots generated using Bray Curtis dissimilarities did not reveal clusters of biofilm bacterial communities at the family level (Fig. 4) as a function of water supply type, water source, physiography, pipe material, well age, well visibility aboveground, disinfection method, sampling time of day, sampling season, and whether or not the aerator was removed. This was confirmed using pairwise PERMANOVA analyses of the community dissimilarity matrix for each variable with a Bonferroni correction for multiple comparisons, which found all $P > 0.058$ (pairwise PERMANOVA). Previous work with simulated and operating water distribution systems, as well as with simulated and actual premise plumbing, has suggested that source water chemistry, geographic location, available carbon, disinfectant type and residual, stagnation period, water age, water heater temperature, and pipe material, and interactions between these factors, influence microbiome composition of biofilms formed on plumbing surfaces (Falkinham, 2011; Ji et al., 2015; Proctor et al., 2017; Wang et al., 2014; Xue et al., 2020).
FIG 4 nMDS plot using Bray Curtis dissimilarity matrix at the family level (N = 22,000 sequences per sample) (stress = 0.18) as a function of (a) pipe material and disinfection type and (b) physiography and well age (public supply samples marked as such rather than by physiography).

3.4.2. LEfSe Results

Linear discriminant analysis (LDA) effect size (LEfSe) was performed on all 29 sequenced biofilm samples (18 replicate samples from nine homes, 11 samples from individual homes) to determine potential antagonists or mutualists of NTM by identifying biomarkers of sample groups with and without myco16S genes. Four prokaryotic families under the phyla Proteobacteria, Firmicutes, and Bacteroidetes were identified as biomarkers for the absence of myco16S (Table S6). These families are comprised of either environmental or gut bacteria, and three out of four contain known human pathogens. For example, the Pseudomonadaceae family has members that are often found in soils and water and contains the species Pseudomonas aeruginosa, a known opportunistic pathogen (Brennan and Geddes, 2002; Palleroni, 1981). The others are enteric bacterial families, including the currently assumed nonpathogenic Rikenellaceae family (Dworkin and Falkow, 2006), the Erysipelotrichaceae family (Kaakoush, 2015), and the unnamed
Bacillales Family XI that contains some opportunistic pathogens, such as *Gemella morbillorum* (Vos et al., 2011). These findings suggest that mycobacteria are less likely to live with other bacteria that prefer complex ecosystems with competition, as is found in the human gut microbiome. This aligns with the existing understanding of mycobacteria as a group that thrives with less competition (Falkinham, 2002). Identifying such markers of the absence of NTM could be a starting point for ecological engineering of premise plumbing microbiomes by exploiting antagonistic relationships as suggested by Wang et al. (2013a), but further studies are required to evaluate this possibility. Additionally, as Wang et al. mention, and as our findings suggest, the search for potential antagonists to ecologically engineer mycobacteria out of drinking water systems may be complicated by the growth of other opportunistic pathogens, such as those that belong to the other taxonomic families identified as biomarkers for the absence of myco16S.

For the presence of myco16S, nine prokaryotic families in addition to the expected *Mycobacteriaceae* family were identified. Among them was *Beijerinckia* family, a family of carbon cycling and dinitrogen fixing bacteria, as previously described (Dedysh et al., 2016; Singleton et al., 2018). Interestingly, previous studies have reported that the presence of *Methylobacterium*, the genus of the *Beijerinckia* family most frequently detected in our study, corresponded to lower mycobacteria biofilm formation (Falkinham et al., 2016; García-Coca et al., 2020). In contrast to this and in agreement with our findings, other work has suggested a positive association between *Mycobacterium* and *Methylobacterium* in the biofilms of chloraminated systems (Waak et al., 2019b). Thus, the relationship between these genera remains unresolved.

LDA LEfSe was also used to identify biomarkers that distinguished between biofilm samples from homes with private well water versus public water supply (Table S6). From this, five prokaryotic families emerged as biomarkers for samples from homes with private wells and two for homes with public supply. For private well samples, the biomarkers were generally families of typical environmental bacteria, including *Rhodocyclaceae* and *Gemmatimonadaceae* (Corteselli
et al., 2017; Wang et al., 2019). For public supply, the identified biomarkers were *Enterobacteriaceae* and *Muribaculaceae* (previously S24-7), bacterial families common to animal gut microbiomes (Wiesenborn et al., 2019), likely indicative of the surface water source water for most of the public supply samples, and unlikely to have been viable in the biofilms at the sampling time, given the samples were from systems with residual disinfectant (Korzeniewska and Harnisz, 2012). See supplement S2 for more information regarding private well and public supply biomarkers. Due to the lack of existing information about private well household plumbing biofilms, no direct comparisons can be made between these findings and those of other studies.

4. Conclusion

Mycobacteria were detected using qPCR and amplicon sequencing in the household plumbing biofilms of homes with private wells across physiographic regions in NJ. Biofilms from private well households, including many without disinfection selection pressures, were no less likely than those from NTM patient public supply households to contain NTM. It is recommended that future studies of NTM in biofilms of homes served by private wells include cultivation-, viability-, or activity-based biomolecular techniques to confirm that these observations correspond to viable NTM. RF regressions using the survey and environmental data collected in this study suggested that physiography and indoor environments or use patterns may influence mycobacteria abundance in home biofilms in homes with private wells. However, the models did not explain more than 21% of the variation in NTM observed, which may indicate that a larger sample size and/or measurement of factors not included here, such as water age or flow velocity, is needed. This study was limited by the need to stop sampling in participant homes as a result of the COVID-19 pandemic. Collection of more samples is recommended when possible to facilitate more robust comparisons. While *M. avium* was not detected in any samples, amplicon sequencing results indicate potential presence of pathogenic mycobacteria species, *M.*
abscessus and M. xenopi, the confirmation of whose presence can be the focus of future investigations. Through the LEfSe biomarker analysis, one microbial family that does not contain known human pathogens emerged as a marker for the absence of mycobacteria Rikenellaceae. Understanding such ecological relationships may be useful for engineering healthy plumbing microbiomes. As the burden of disease from NTM continues to rise, and as 43 million Americans use largely disinfection-free private wells, future work should consider levels of microbial contaminants beyond fecal indicators, including NTM, in private well and public systems, alike.

5. Acknowledgements

Thanks to Cassidy Schneider and Rachel Feldstein for assisting in field sampling. Thanks to Melissa Tu for assisting with qPCR. Thanks to Thomas Scanlin and Sugeet Jagpal for help in recruiting participants. Thanks to Alessia Eramo and Aishwarya Deshpande for assistance with qPCR and DNA extraction methods.

6. Author attributions & citation

Conception/ design of the work: CV, NLF, MLG. Data collection: SB, DP. Data analysis & interpretation: SB, NLF. Drafting the article: SB. Critical revision of the article: SB, CV, MLG, NLF.

III. POTENTIAL FOR NONTUBERCULOUS MYCOBACTERIA PROLIFERATION IN NATURAL AND ENGINEERED WATER SYSTEMS DUE TO CLIMATE CHANGE: A CRITICAL LITERATURE REVIEW

1. Introduction

1.1. Disease prevalence and impact

Infections caused by nontuberculous mycobacteria (NTM) have gained increasing attention over the past several decades in medical and engineering fields due to their high cost (Collier et al., 2021), difficult treatment (Daley et al., 2020), and increasing prevalence (Ratnatunga et al., 2020). NTM, otherwise referred to as “atypical mycobacteria,” “environmental mycobacteria,” or mycobacterial species other than Mycobacterium tuberculosis and Mycobacterium leprae, are ubiquitous in water and soil environments and include many opportunistic pathogens such as those in the Mycobacterium avium complex (MAC) (Falkinham, 2016). Exposures to NTM species by immunocompromised individuals can result in a variety of infection and disease outcomes including pulmonary disease, as well as extrapulmonary and disseminated diseases (Daley et al., 2020; Falkinham, 1996; Wolinsky, 1979). Cases of NTM infections and disease come with a high price tag. A US Centers for Disease Control (US CDC) study estimated that NTM infections were the costliest primarily water-based infections, costing $1.53 billion in direct healthcare costs in the US in 2014 (Collier et al., 2021). Due to the increasing instances of disease (Donohue, 2018) and prevalence in drinking water systems (Falkinham, 2015b), M. avium has been included in all four US Environmental Protection Agency (USEPA) Drinking Water Contaminant Candidate Lists (CCL-4) (USEPA, 2016). While concerns about NTM disease typically focus on immunocompromised individuals who comprise a large fraction of patients, researchers have recently noted that NTM disease prevalence is increasing in immunocompetent populations, raising the public health concern around these bacteria (Stout et al., 2016).
1.2. The nexus of pathogen, climate, and human interactions- a roadmap to risk analysis

The majority of NTM infections are believed to be acquired through human-environment interactions, such as by inhalation of infected particles and bioaerosols (Falkinham, 2013). NTM have been identified in soils (De Groote et al., 2006; Kirschner et al., 1992), household dust (Dawson, 1971), freshwater sources (Roguet et al., 2018), public water supplies (Kotlarz et al., 2018; Waak et al., 2019a) and in plumbing biofilms in homes using public and private well water supplies (Blanc et al., In revision; Feazel et al., 2009; Haig et al., 2020; Wang et al., 2012b; Xue et al., 2020). As environmental bacteria, NTM can be expected to be influenced by changing climatic conditions. Since the dawn of the industrial revolution, emissions of greenhouse gases such as carbon dioxide (CO$_2$) and methane (CH$_4$) have increased, leading to an accelerated increase in average global land and ocean temperatures at a rate of approximately 0.2°C per decade (Allen et al., 2019). Increases in atmospheric carbon dioxide and global temperatures have rippling effects (IPCC, 2014). Some of these include increases in humidity, intense precipitation and drought, changes in sea level, increases in the frequency and/or intensity of natural disasters, and extended warm seasons. These changes lead to numerous secondary, or cascading effects such as increased flooding, saltwater intrusion, mold exposure, and water and food instability (Allen et al., 2019). Coastal regions are particularly affected by climate change, and in the US, these are many of the same regions (e.g., Hawaii, Louisiana, California, Florida) with the majority of NTM infections (Adjemian et al., 2012; De Groote and Huitt, 2006; Strollo et al., 2015). Researchers from the Intergovernmental Panel on Climate Change (IPCC) suggest that the trajectory of global emissions (Allen et al., 2019) and the associated societal and population dynamics (O'Neil et al., 2017; Riahi et al., 2017) are highly likely to influence the magnitude of the abovementioned effects.
This analysis seeks to understand how climate change might influence NTM in water environments, natural and engineered, and how that influence might translate to future risk of NTM exposure and infection.

As cases of NTM disease continue to climb, global climate change accelerates, and population dynamics shift, there is an urgent need to understand how risk of NTM infections might be expected to change in the coming decades. Literature reviews have been published that synthesize information about NTM epidemiology and ecology (Falkinham, 2010; Honda et al., 2018; Nishiuchi et al., 2017; Parikh et al., 2019; Pereira et al., 2020), and specifically about NTM and natural disasters (Honda et al., 2015). However, there remains a gap in the literature as to how the combination of changing climates and human dynamics might influence NTM risk, despite evidence that climate (Chou et al., 2014; Thomson et al., 2020) and societal (Lipner et al., 2017) factors can influence NTM disease, and that these factors influence each other (Riahi et al., 2017). Therefore, the overarching objective of this review is to synthesize the wealth of information about NTM, climate, and human system dynamics to characterize the processes that may influence future human risk to NTM infections (Fig. 5). The specific goals of this review are to capture the primary interactions occurring at the interface of several systems, to highlight the gaps in understanding how the systems function and interact, and to identify potential scenarios that might lead to different levels of future risk. This analysis is intended to support future model-
based risk analyses of NTM, as well as to inform decision making by water managers to mitigate potential risk.

2. Methods

This analysis was completed using a literature review in January 2021 using Google Scholar, Engineering Village, Web of Science, and PubMed databases. Search terms such as “fate and transport,” “decay,” “growth,” “dynamics,” “temperature,” “precipitation,” “flood,” “drought,” “saltwater,” “nutrient,” “disaster,” “climate change,” “risk,” “QMRA,” and “nontuberculous mycobacteria” were used in several databases to obtain a base set of studies. Relevant studies were categorized into groups including pathogen dynamics, ecology, climate change, drinking water systems, and water reuse systems. Once organized, studies were evaluated in closer detail, and a reverse search was conducted. As information pertaining to system dynamics was noted, connections were made that inform the analysis described in sections 3.1–3.3. Although several hundred studies were considered, only the most relevant have been included in this analysis. As these studies were evaluated, information relating to NTM quantities measured were also noted to inform future work involving quantitative microbial risk analysis (QMRA) for NTM species. Discussion sections 3.4–3.5 were informed by a smaller number of more specific studies. The included studies covered a broad range of topics, including NTM survival mechanisms, environmental influences on NTM prevalence and abundance, the increasing nature of NTM diseases, and climate and engineering influences on NTM in various environments.

3. Results and Discussion

3.1. Biological, ecological, and physiological characteristics

To hypothesize how risk of NTM infections might change under future circumstances, some key characteristics that affect the survival of NTM must be understood. First, NTM have an outer membrane rich in lipids and composed of mycolic acids that provide them with an array of survival-enhancing qualities including hydrophobicity, impermeability, and slow growth.
(Brennan and Nikaido, 1995). Because of their hydrophobicity, NTM tend to form biofilms on surfaces such as pipe walls or soil particles (Schulze-Röbbecke et al., 1992), to which they can also adsorb due to their negative surface charge (Bendinger et al., 1993; Brooks, 1983; Lytle et al., 2004). NTM can survive and replicate within several protozoan species (Strahl et al., 2001), primarily free-living amoeba in both the trophozoite and cyst phases (Barker and Brown, 1994; Cirillo et al., 1997; Delafont et al., 2016; Prasad and Gupta, 1978; Thomas et al., 2008). Some NTM species can adapt by sharing genetic information through plasmid-mediated horizontal gene transfers (Nguyen et al., 2010). Biofilms provide an ideal habitat for NTM to share information about antibiotic resistance genes (ARGs) as well as metal resistance genes (MRGs) (Kimbell et al., 2020; Liu et al., 2016). NTM also have other adaptive capabilities that aid in survival, such as the ability to enter a stationary phase of reduced cell activity under starvation (Smeulders et al., 1999) and acid stresses (Bodmer et al., 2000). They are known to survive typically extreme environments, such as acidic (Bodmer et al., 2000), warm (Schulze-Röbbecke and Buchholtz, 1992), oligotrophic (Smeulders et al., 1999; Zhu et al., 2019), and microaerobic environments (Falkinham, 2013; Kirschner et al., 1992).

Environmental concentrations of NTM species vary depending on the location, measurement method, and sample matrix (Table 5). As an example, both the Finland brook study (Iivanainen et al., 1993) and the southeastern US swamp study (Kirschner et al., 1992) cultivated mycobacteria from water and found average values that differed by two orders of magnitude, owing to different environmental conditions. Moreover, measurements for abundance of mycobacteria in aerosols are not readily comparable to those for water or biofilms, even in the same study, due to differences in measurement units (Kirschner et al., 1992). For these reasons, future investigations that seek to calculate quantitative risks analyses of NTM should consider the specific location, source, predicted transmission route, and species of interest for the exposure assessment. Some researchers have already attempted QMRA for NTM species or have investigated dose-response
relationships, and future work may consider building off of those studies (Breuninger and Weir, 2015; Cui et al., 2017; Hamilton et al., 2017a; Hamilton et al., 2017b; Rice et al., 2005).

**TABLE 5.** Measurements of mycobacteria across several studies spanning decades, geographics locations, sample types, and measurement methods. Note, this table is not exhaustive.

<table>
<thead>
<tr>
<th>Data measured</th>
<th>Value</th>
<th>Unit</th>
<th>Matrix</th>
<th>Sample origin</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture counts</td>
<td>Range: 0 – 10⁶</td>
<td>CFU/cm²</td>
<td>Biofilm</td>
<td>DWTP and home plumbing, Germany and France</td>
<td>(Schulze-Röbbecke et al., 1992)</td>
</tr>
<tr>
<td></td>
<td>x̄ = 4.06 × 10³</td>
<td></td>
<td></td>
<td>Peatland influenced brook, Finland</td>
<td>(Ilvanainen et al., 1993)</td>
</tr>
<tr>
<td></td>
<td>σ = 1.49 × 10⁶</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range: 10 – 2200</td>
<td>CFU/L</td>
<td>Water</td>
<td>DWTP a, Spain</td>
<td>(Corsaro et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>x̄ = 618</td>
<td></td>
<td></td>
<td>DWTP a, Spain</td>
<td>(Corsaro et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>σ = 603</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range: 0 – 4.8 × 10⁴</td>
<td>CFU/L</td>
<td>Water</td>
<td>DWTP a, Spain</td>
<td>(Corsaro et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>x̄ = 2.11 × 10⁴</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>σ = 1.83 × 10⁴</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range: 0 – 1.65</td>
<td>CFU/cm² /hr</td>
<td>Ejected</td>
<td>DWTP a, Spain</td>
<td>(Corsaro et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>x̄ = 0.82</td>
<td></td>
<td>droplets</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>σ = 0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range: 0 – 66.1</td>
<td>CFU/m³/hr</td>
<td>Aerosol</td>
<td>DWTP a, Spain</td>
<td>(Corsaro et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>x̄ = 24.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>σ = 26.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culture Range</td>
<td>1 – 50 (78)</td>
<td>CFU/L</td>
<td>Water</td>
<td>DWDS c, France</td>
<td>(Le Dantec et al., 2002)</td>
</tr>
<tr>
<td></td>
<td>51 – 500 (21)</td>
<td>(% of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt; 500 (1)</td>
<td>samples)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culture P/A</td>
<td>4/16 (25)</td>
<td>P/A (%)</td>
<td>Biofilm</td>
<td>DWTP a, Spain</td>
<td>(Corsaro et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>46/165 (28)</td>
<td></td>
<td></td>
<td></td>
<td>(Falkinham, 2011)</td>
</tr>
<tr>
<td></td>
<td>4/55 (7)</td>
<td></td>
<td></td>
<td></td>
<td>(Falkinham, 2011)</td>
</tr>
<tr>
<td></td>
<td>2/10 (20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5/16 (31)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5/12 (42)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>41/60 (68)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>32/89 (36)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>33/48 (69)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30/36 (83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>90/142 (63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qPCR</td>
<td>Range: &lt; 4.5 × 10² – 2.4 × 10⁴</td>
<td>copies</td>
<td>Biofilm</td>
<td>No disinfectant residual</td>
<td>(Waak et al., 2019b)</td>
</tr>
<tr>
<td></td>
<td>x̄: &lt; quantitation</td>
<td>atpE gene/cm²</td>
<td></td>
<td>DWDS c, SW d, US</td>
<td></td>
</tr>
</tbody>
</table>

a DWTP: Drinking Water Treatment Plant
b DWDS: Drinking Water Distribution System
c SW: Surface Water
d GW: Ground Water

h qPCR: Quantitative Polymerase Chain Reaction

Note: This table is not exhaustive.
<table>
<thead>
<tr>
<th>Range: $6.0 \times 10^2$ – $4.8 \times 10^6$</th>
<th>copies $atpE$ gene/cm$^2$</th>
<th>Biofilm Chloraminated DWDS $^c$, SW $^d$, US</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{x}: 9.3 \times 10^4$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Range: $< 200 – 7.76 \times 10^6$ | copies $atpE$ gene/L | Water Chloraminated DWDS $^c$–$^e$ mixed, US |
| Range: $2.0 \times 10^4$ – $1.3 \times 10^7$ | myco16$^f$ gene copies/L | No disinfectant residual DWDS $^c$, mixed, Netherlands |

| Range: $8.6 \times 10^2$ – $4.4 \times 10^7$ | myco16$^f$ gene copies/L | Water SWSS $^g$, China |

| Range: $10^4$ – $10^7$ | myco16$^f$ gene copies/L | Water Chlorinated DWDS point-of-use, mixed, China |
| $\bar{x} = 10^5.78$ | $\sigma = 10^{0.72}$ |  |
| Range: $6.7 \times 10^3$ – $1.9 \times 10^8$ | copies $atpE$ gene/L | Water Lakes, France |
| $\bar{x} = 2.16 \times 10^5$ |  |

$^a$ DWTP = drinking water treatment plant

$^b$ focused on $M. avium$, $M. intracellular$, and $M. scrofulaceum$

$^c$ DWDS = drinking water distribution system

$^d$ DW = drinking water, SW = surface water, GW = groundwater; if SW or GW not listed, information was not available in referenced paper

$^e$ samples from premise plumbing in homes connected to the DWDS

$^f$ myco16 refers to the *Mycobacterium* spp. -specific fragment of the 16S rRNA gene

$^g$ SWSS = secondary water supply system, which refers to water storage and pressure systems connected to larger buildings, such as rooftop tanks

$^h$ Evaluated for but did not find $M. avium$-specific genes in any samples
3.2. Climate change considerations

3.2.1. Temperature

For NTM in surface water or shallow groundwater environments (Taylor and Stefan, 2009), warmer temperatures could provide more favorable conditions given that NTM species better survive and grow in warm waters than in cold (George et al., 1980; Torvinen et al., 2007). For example, NTM were found to be more prevalent in southeastern US waters as compared to waters in the northeastern US, an observation that researchers attributed to the relatively longer period of time that water in the southeast spends above 15.5°C and the smaller time period that it spends below 9.4°C (George et al., 1980). Similarly, MAC species were found in greater abundance in warm water (Dailloux et al., 1999), such as surface water during warm seasons (Kirschner et al., 1992). Correlations were observed between temperature and mycobacteria detected by qPCR in a coastal lagoon (Jacobs et al., 2009).

During increasingly frequent extreme heat events (Allen et al., 2019), NTM species may be uniquely adaptable. Some species (e.g., *M. xenopi*) have been shown to survive in extremely warm waters, at temperatures as high as 60°C (Schulze-Röbbecke and Buchholtz, 1992). While disease risk can be greater in warmer regions (De Groote and Huitt, 2006; Honda et al., 2015; Honda et al., 2018), a 16 year study in Australia found no such association (Thomson et al., 2020), potentially indicating that the relationship between temperature and disease risk may be complicated by additional factors.

As home plumbing systems are considered transmission routes, one might consider how changing temperatures influence NTM in water once it is treated and within a distribution system. Several studies have reported seasonal increases in NTM in drinking water distribution systems during warm times of the year (Liu et al., 2019; Perez-Martinez et al., 2013; Thomson et al., 2013b; van der Wielen and van der Kooij, 2013) and correlations between water temperature and mycobacteria abundance (Ling et al., 2018; Revetta et al.,
suggesting that increased temperature carries potential to increase NTM abundance in these engineered systems as it provides a more suitable environment for survival (Liu et al., 2019). Considering climate change, Walker et al. suggest that longer periods of warmer temperatures within buildings could result in increased regrowth in pipes as building operators struggle to maintain in-pipe water temperatures below 20°C (Walker, 2018). This has already been observed as an issue in the United Kingdom in the context of Legionella (Walker, 2018), which is sensitive to disinfectant residual that degrades at warmer temperatures (Sebakova et al., 2008). However, NTM have been found to resist or be selected for by disinfection (Falkinham, 2009; Steed and Falkinham, 2006), so the impact that disinfectant degradation may have on NTM remains uncertain. Further, as stated in Section 3.1, NTM can live within amoeba, some of which (e.g., Vermamoeba) have been found in greater abundance in drinking water distribution systems during warmer seasons (Delafont et al., 2016), indicating potential for NTM to enhance survival by living within the amoebas that also prefer warmer conditions.

3.2.2. Humidity

In humid environments, NTM may be more likely to survive and transport in aerosols that remain suspended for longer periods of time. With increased temperatures, increases in humidity have been observed and are projected to continue in already humid regions of the world (Barreca, 2012). A study by Lin and Marr (2019) found increased survival with relative humidity for bacteria in aerosols, including M. smegmatis. Aerosol droplets can resist full evaporation, retaining liquid water for up to an hour when the relative humidity is higher than 55%. This suggests that if NTM aerosolize in humid environments, these microorganisms may have ample time to transport through air in the remaining liquid (Lin and Marr, 2019).
3.2.3. Sea level rise

Rising temperatures also have important implications for NTM due to the secondary effect of sea level rise, influencing salinity gradients in estuaries and the saltwater/freshwater interface in aquifers (IPCC, 2014). Research suggests that most NTM species appear in greater abundance in waters with lower salinity (Dailloux et al., 1999), and that salinity above 3 g/L can inhibit replication of most NTM species (George et al., 1980). In a coastal lagoon study, a negative correlation between salinity and mycobacteria gene abundance was one of three factors (along with total nitrogen and dissolved oxygen) used to build a model to predict measured values with 83% agreement, suggesting that the negative association is important to understanding mycobacteria abundance in waters with a salinity gradient (Jacobs et al., 2009). During storms with winds and heavy precipitation, mixing of saltwater and freshwater may result in greater aerosol production efficiency (Honda et al., 2015; Parker et al., 1983). As the salinity gradient migrates inland, this could result in more aerosols forming closer to the land where more people may be exposed via inhalation. Therefore, although the abundance of NTM species in mixed saltwater/freshwater zones may be expected to generally decrease as the salinity gradient migrates inland, the potential for greater aerosolization efficiency and exposure complicates risk predictions.

3.2.4. Precipitation, hydrology, and water quality

Increases in heavy precipitation are projected to lead to increases in flooding and runoff (Allen et al., 2019) that can elevate turbidity and nutrient concentrations, and decrease oxygen content in receiving waters (IPCC, 2014). Due to their hydrophobic membranes that enable attachment to particle surfaces (Bendinger et al., 1993; Brooks et al., 1984), increased turbidity may lead to increased abundance of NTM in receiving surface waters. Previous work supports this positive association between NTM and turbidity (Falkinham et al., 2001; Jacobs et al., 2009; Vaerewijck et al., 2005). NTM have also been associated with various
metals, such as nickel, chromium, and iron, potentially for similar mechanistic reasons (Robinson, 2019), further suggesting potential for NTM concentrations in receiving water to increase with increased runoff carrying metal pollution. However, this could be complicated by increased flushing rates, which may reduce concentrations of the abovementioned pollutants, underscoring the importance of local models and understanding (IPCC, 2014). Additionally, the increase in nutrients as a result of increased runoff could create eutrophic conditions that have previously been correlated with mycobacteria gene concentrations (Jacobs et al., 2009). This could be because of the capacity of NTM to survive under anoxic or anaerobic stresses that other organisms cannot (Pereira et al., 2020), as well as their preference for warm water associated with eutrophic conditions (IPCC, 2014). Although NTM are often considered oligotrophic (Falkinham, 2009), they can adapt to nearly opposite environmental conditions imposed by increased runoff (Jacobs et al., 2009). This indicates that increased precipitation intensity resulting in eutrophication may increase NTM in receiving surface water environments. Combined sewer overflows (CSOs) are also a concern with respect to increased runoff, as stormwater and wastewater flow together and carry pathogens directly to receiving water bodies (Patz et al., 2008). Further, periods of intense precipitation have the potential to create sanitary sewer overflows, resulting in more concentrated waterbody contamination with pathogens including NTM (Nasrin et al., 2017).

Increases in precipitation intensity also carry the potential for increased flooding and associated impacts to shallow, unconfined groundwater (IPCC, 2014). Similar to the discussion above, increased flooding could result in increased transport of contaminants (e.g., nutrients, organic matter, metals, microorganisms) that can introduce the contaminants to shallow groundwater as the subsurface becomes saturated (Earman and Dettinger, 2011; Kløve et al., 2014). Because NTM adhere to surfaces (Bendinger et al., 1993; Bolster et al., 2009), there is uncertainty about how efficiently they may be removed by soils during
infiltration. Flooding could also lead to waterlogging, blocking off air flow between the atmosphere and subsurface, creating anoxic conditions (Estop-Aragonés et al., 2012; Kløve et al., 2014) for which NTM are well suited (Falkinham, 2013; Pereira et al., 2020). Alternatively, in other situations, flooding could introduce oxygenated water into the water table. However, oxidation of associated nutrients and organic matter could consume the oxygen nonetheless, again resulting in reduced oxygen conditions (Yu et al., 2015). Flooding of typically unsaturated soils also may release naturally accumulated salts into aquifers, potentially increasing salinity (Earman and Dettinger, 2011), a factor that could negatively influence NTM survival (Jacobs et al., 2009), as described above.

In regions where precipitation is projected to decrease or flow to be seasonally reduced, water quality and NTM considerations differ from those discussed for increasingly wetter conditions. When surface water flow decreases as a result of drought or reduced snowpack, the volume of water available to dilute pollutants (e.g., nutrients, organic matter, microorganisms) (Weinrich et al., 2010) introduced to surface water by wastewater decreases (IPCC, 2014). Cultivable NTM (Jjemba et al., 2010) and mycobacteria genes (Amha et al., 2017; Chen et al., 2019) have been found in wastewater effluents. Therefore, low flows, exacerbated by potential increases in evapotranspiration rates, could allow for increased concentrations of NTM in water sources simply by way of less dilution of wastewater effluents. In regions with high rates of de facto water reuse, therefore, this could degrade source water quality with respect to NTM. In areas with combined sewers, low flow as a result of drought would likely accompany a decrease in combined sewer overflows, decreasing associated microbial loads in connected water bodies. However, in areas with low flow due to earlier and reduced snow melt, combined sewer overflows may still occur with precipitation events, and would be met with less water available for dilution, increasing microbial concentrations (Jalliffier-Verne et al., 2015). In groundwater, where reduced
precipitation and increased evapotranspiration also decrease flow and lower the water table, there are changes in the hydraulic gradient where aquifers meet surface water (Kløve et al., 2014; Treidel et al., 2011). In coastal areas, this could lead to increased saltwater intrusion, as has been seen in areas of California (Ashbolt, 2010), expected to negatively influence the suitability of the habitat for NTM (Kløve et al., 2014). At the headwaters, decreased baseflow from cooler groundwater into streams could result in increased surface water temperature (Earman and Dettinger, 2011), favorable for NTM (Kirschner et al., 1992). Decreased flow in aquifers could also lead to increased residence time and, therefore, contact time between subsurface rock and water, increasing the opportunity for minerals to leach (Earman and Dettinger, 2011). Minerals related to rocks such as hematite and goethite were previously associated with NTM abundance (Robinson, 2019). Therefore, decreased groundwater flow has the potential to increase or decrease NTM abundance, depending on whether salinity or mineral concentrations control their survival and growth.

3.2.5. Natural disasters

With tropical cyclones, there is potential for increased aerosolization of NTM as freshwater, saltwater, and soils mix, as well as for transport of the bioaerosols with cyclone winds (Honda et al., 2015). According to the IPCC, it is likely that the average tropical cyclone wind speed and precipitation will increase in coming decades (Field et al., 2012). Thus, as wind intensity of cyclones is projected to increase with climate change (Field et al., 2012), so too might NTM exposure through bioaerosols. A recent study of climatic factors and NTM disease found associations that supported this idea. Associations between heavy rainfall and disease differed in direction between humid, tropical storm environments and arid environments: risk in humid environments increased and risk in arid environments decreased, both with several months lag time (Thomson et al., 2020). The researchers suggested that this may be because the humid environments were subject to tropical storms with strong winds.
capable of transporting aerosolized NTM far distances. In addition, Honda et al. (2015) suggested that the relationship between NTM and amoeba may further contribute to infection risk post-disaster, as NTM within amoeba may better survive when displaced from originating water sources by cyclones. Previous investigations of spatial clusters of NTM disease have found hot spots in regions with frequent disasters, such as in Florida, Hawaii, and Louisiana, suggesting that increased exposure to extreme events may relate to risk (Adjemian et al., 2012). Additional details on the potential relationships between NTM and natural disasters can be found in a review by Honda et al. (2015). These authors also acknowledge that, to fully assess risk in this context, additional factors such as changes in host susceptibility with trauma and food/water insecurity need to be considered and better understood as NTM are primarily opportunistic pathogens and may find more suitable hosts during disaster periods (Honda et al., 2015).

Changes in engineering and water management practices

Meeting water demands while adapting to challenges including reduced snowpack, unpredictable rainfall, and saltwater intrusion, as well as potential increased water requirements for energy production, will require flexibility and creativity (Ashbolt, 2010). Adaptations to these stresses, as with any water management changes, will influence the prevalence and abundance of contaminants, including NTM. Some practical adaptation choices to address the challenges of water scarcity and water quality are discussed in this section as specifically related to NTM risk.

3.2.6. Managing water scarcity with reclaimed water

As water resources become more stressed and water reclamation trends continue upward (Ghernaout, 2018), there are several considerations to reduce the risk of exposure to NTM. Water reclamation or reuse includes wastewater reuse, as well as reclamation of stormwater for uses such as washing, irrigating, cooling, toilet flushing, or even drinking (Ghernaout et
Wastewater reuse today occurs both intentionally through extensive treatment and redistribution (Amha et al., 2017; Böckelmann et al., 2009) and non-intentionally (de facto reuse) by way of poor sanitation (Huo et al., 2021) and low flow rivers that accept wastewater effluents while feeding drinking water influents (Rice et al., 2005).

To reduce NTM in reclaimed wastewater, several studies suggest that some current practices may be insufficient (Amha et al., 2017; Jjemba et al., 2014). For example, treatment including biological reactors, microfiltration, and chlorination selected for mycobacteria in wastewater effluents (Chen et al., 2019). In another study, when conventional treatments, membrane bioreactors (MBR), and disinfection reduced NTM at the wastewater treatment plant, they regrew to densities 10 times those of fecal indicators along the redistribution system, similar to behavior in potable water systems (Jjemba et al., 2010). Based on existing knowledge about NTM’s resistance to disinfection and tendencies for regrowth (Falkinham, 2009; Li et al., 2016; Oriani et al., 2018; Wang et al., 2012b; Xu et al., 2017), these observations are not surprising. In contrast, studies suggest that a treatment train including conventional treatments in addition to reverse osmosis (RO) may be effective at removing NTM, other pathogens, and even ARGs that contribute to their virulence (Böckelmann et al., 2009; Harb et al., 2019; Stamps et al., 2018). This could be because RO exploits the negative charge of NTM (Lytle et al., 2004) to exclude the cells (Yangali-Quintanilla et al., 2011).

Challenges to this type of treatment, however, do exist. The high cost, energy intensity, production of RO concentrates, and potential for dangerous disinfection byproducts formed by advanced oxidation processes (AOPs) that typically follow RO are all practical concerns (Roccaro, 2018). These challenges could be overcome by implementation of alternative treatments, such as nanofiltration, that similarly take advantage of NTMs’ negative charge to exclude them from finished water but with less drawbacks in terms of energy usage and cost (Yangali-Quintanilla et al., 2011). Disinfection by UV rather than chemical disinfectants
could also be employed, though the effectiveness of UV varies by NTM species (Lee et al., 2010). Further, Falkinham (2009) highlights that this may result in mutations that require investigation. Temperature management could also be used to control NTM. Treating with heat above 53°C has been found to control MAC species, however, other species (e.g., *M. xenopi*) resist temperatures above that (Norton et al., 2004).

Even if treatments successfully reduce NTM concentrations in reclaimed water at the treatment plant, potential for regrowth along redistribution systems remains a concern. One reason for this is thought to be availability of nutrients in reclaimed water. Although NTM are often considered to be oligotrophic (Falkinham, 2009) and studies have shown that assimilable organic carbon (AOC) may not be a limiting nutrient for all NTM (Liu et al., 2019; Zhu et al., 2019), others suggest that severely limiting AOC (< 10 μg/L) may inhibit growth of the most clinically relevant and likely to be regulated MAC species (Norton et al., 2004; van der Wielen and van der Kooij, 2013). Studies of reclaimed water treated conventionally and with MBR have found relatively high AOC concentrations (range 45 – 3200 μg/L, median 450 μg/L), suggesting that additional processes would need to be undertaken to drastically lower the concentrations (Weinrich et al., 2010). Chemical disinfectants increase AOC concentrations (Liu et al., 2015; Weinrich et al., 2010). Thus, controlling microbial growth by engineering low nutrient conditions rather than by using disinfectants, as is done in the Netherlands (van der Wielen and van der Kooij, 2013), would likely reduce AOC throughout redistribution (Garner et al., 2016). Reducing corrosion in the redistribution system may also reduce the likelihood of NTM proliferation by controlling the surface area available for attachment and preventing additional nutrients from entering the system (Norton et al., 2004). Corrosion control could also reduce the virulence of the surviving NTM by limiting horizontal gene transfer of ARGs and MRGs (Kimbell et al., 2020; Liu et al., 2016; Pal et al., 2015). However, corrosion control mechanisms should be
thoughtfully considered because research has also shown that phosphates, often used for corrosion control, may aid the growth of NTM in distribution system biofilms (Fang et al., 2009; Garner et al., 2018; Zhu et al., 2019). Another challenge that is particularly relevant when considering reclaimed water distribution in the context of climate change is maintaining a cool temperature to limit growth activity of NTM in biofilms (Garner et al., 2018). While this remains a challenge, potential solutions such as providing a minimum pipe to subsurface depth or distance from electric cables, or providing shade or vegetation cover above distribution pipes can be investigated for efficacy and practicality (Agudelo-Vera et al., 2020).

In addition to using reclaimed wastewater, rainwater reclamation is another strategy to make use of finite water resources in areas facing unreliable water supply. This can be done through rainwater catchment systems, such as roof-harvested rainwater systems that direct rainwater from a building’s roof into a barrel, tank, or collection channel for uses including washing, toilet flushing, showering, irrigating, or drinking (Hamilton et al., 2017a). The water quality of roof-harvested rainwater has been investigated for risk of exposure to opportunistic pathogens, including NTM. Some studies have found NTM species, including \textit{M. avium} and \textit{M. intracellular}, frequently in roof-harvested rainwater (Hamilton et al., 2017a; Hamilton et al., 2018), while others in different locations found them seldom or did not find them at all (Albrechtsen, 2002; Zhang et al., 2020). A comprehensive QMRA for MAC species in roof-harvested rainwater resulted in the conclusion that use of roof-harvested rainwater can be low risk if used for car and clothing washing, or toilet flushing, but that it should not be used for drinking, irrigating, or showering (Hamilton et al., 2017a). Interestingly, although much concern around NTM involves inhalation, this QMRA study found the inhalation risk to be lower than the risk from ingestion. Importantly, however, the QMRA also suggested that the cumulative risk reflective of solely relying on roof-harvested rainwater was higher than
acceptable (10^4), suggesting that risk of infection could increase if no other water sources were available (Hamilton et al., 2017a). Thus, as water resources become more scarce in some regions, water managers may consider ways to ensure that residents do not rely solely on untreated roof-harvested rainwater, as is already the case in 13.6% of households in Queensland, Australia (ABS, 2013).

3.2.7. Reclaimed water storage/ artificial aquifer recharge

There is much uncertainty around the potential for proliferation of NTM as a result of managed artificial recharge. Some researchers suggest that storing reclaimed water in a high quality aquifer can improve water quality through natural microbial processes (Dillon, 2009; Dillon et al., 2010), while others suggest that introducing reclaimed water into a clean aquifer may result in unintended consequences, such as distribution of ARGs (Harb et al., 2019). While Harb et al. (2019) suggest that the original groundwater microbiome dictates the final recharged aquifer microbiome, they also report evidence of wastewater influence by way of ARGs, highlighting the uncertainties of the extent of wastewater influence. This could be cause for concern about potential acquisition of ARGs because NTM are capable of horizontal gene transfer (Nguyen et al., 2010). However, both NTM and ARGs have been identified in aquifers without the influence of managed artificial recharge (Le Dantec et al., 2002; Machado and Bordalo, 2014; Unno et al., 2015), and some of the water treatments described in Section 3.3.1 may reduce the likelihood of additional ARGs and NTM entering aquifers. In contrast, unmanaged artificial aquifer recharge via disposal of wastewater or stormwater into aquifers without water quality considerations could have negative effects in terms of NTM and other pathogens (Dillon et al., 2010). For this reason, intentional water reclamation and recharge is a preferred option from a water quality and water quantity standpoint as it provides control and conserves water (Dillon et al., 2019). When considering intentional infiltration of stormwater through wetlands, research has indicated that high
temperatures and nutrient levels can result in NTM presence in the connected aquifer (Jacobs et al., 2009; Tjandraatmadja et al., 2014). Thus, Jacobs et al. (2009) suggest that steps taken to reduce eutrophication in wetlands may also benefit the water quality of aquifers below, which could potentially reduce NTM prevalence.

3.3. Related societal dynamics

Changing societal dynamics feed into climate change, influence water management decisions, and affect host exposure and susceptibility to opportunistic infections. Political will and associated policies that curb climate change by shifting energy and industrial practices will influence the magnitude of the abovementioned climate challenges (O'Neill et al., 2017). Increases in population density and urbanization trends influence air and water temperatures via the urban heat island effect (Rizwan et al., 2008). For example, in Minnesota urbanization increased shallow groundwater temperatures by 3°C compared to temperatures in agricultural areas (Taylor and Stefan, 2009). This increased temperature could provide a more supportive environment for NTM survival and growth (Schulze-Röbecke and Buchholtz, 1992) and is driven in large part by the presence of heat trapping land cover, such as pavements (Taylor and Stefan, 2009). Urbanization and associated impervious land cover combined with increased wastewater loads (associated with increased population density) may also impact water quality through increased nutrients (e.g., nitrogen and phosphorous) from runoff and wastewater, potentially resulting in eutrophication (Huang et al., 2014). Implementation of stormwater best management practices (BMPs), such as rain gardens and bioswales, could offset increases in runoff, but the impact to eutrophication, and ultimately to NTM concentrations, would be complicated by the reduction in flow available for nutrient dilution (Small et al., 2019). Urbanization, often concentrated in coastal and riverine areas, also exposes more people and their water sources to floods and natural disasters (IPCC, 2014), both of which may lead to greater exposure to environmental pathogens, including NTM
(Honda et al., 2015). Population density, a proxy for urbanization, has been associated with NTM disease prevalence in several studies (Lipner et al., 2017; Olivier et al., 2003; Thomson et al., 2020).

**FIG 6** Systems affecting NTM fate and transport considering climate and societal changes. Green arrows indicate a direct relationship (+), red arrows indicate an inverse relationship (-), and purple arrows indicate uncertain relationships. Yellow ovals indicate key nodes that affect the system outcome and connect to other nodes that may be points of intervention. The left side of the figure represents societal and engineering dynamics, and the right represents climate/natural environment dynamics. Note, although the connections here are informed by the multitude of studies reviews for this analysis, there are many uncertainties, and some details are omitted for readability.
Increased temperature and population may also increase energy and water demands (Allen et al., 2019). Water requirements for energy and food production (e.g., biofuel crop production) may further stress water resources affected by droughts and rainfall unpredictability, requiring more water reclamation and storage (Allen et al., 2019; Ashbolt, 2010). With proper water management (i.e., conservation, ‘One Water’ integrated cooperative management among stakeholders) and treatment processes, increased reliance on reclamation and artificial aquifer recharge could result in minimal changes or improvements in NTM exposure risk (Hamilton et al., 2017a; Stamps et al., 2018). In the absence of comprehensive management and treatment, increased NTM exposure from reclaimed or unmanaged aquifer recharge is also possible (Amha et al., 2017; Jjemba et al., 2010).

The impacts of climate and societal change also influence individuals and their risk to opportunistic infections. With increased exposure to extreme weather events, individuals experience trauma and food/water insecurity that decrease their immune systems, increasing susceptibility to opportunistic infections including those by NTM (Honda et al., 2015). Disaster events also may lead to mold and associated mycobacteria growth, potentially increasing exposure to NTM (Torvinen et al., 2006). Individual immune response also declines with age, suggesting that as the global population ages, vulnerability to opportunistic NTM infections will increase (Falkinham, 2009; Ratnatunga et al., 2020). Access to equitable healthcare and education also influence infection outcomes by enabling or preventing access to adequate treatment (Lipner et al., 2017; Ratnatunga et al., 2020).

3.4. Potential scenarios for consideration

To synthesize the information discussed thus far, scenarios are provided that characterize potential outcomes of NTM risk in the context of the detailed shared socioeconomic pathways (SSPs) developed by IPCC scientists (O'Neill et al., 2017; Riahi et al., 2017). Fig. 6 shows the system dynamics involved that can be applied to any of the SSPs, following the
direct (green) and inverse (red) relationships. These scenarios represent merely a few of the infinite potential pathways and outcomes of future change. They are a result of assumptions combined with quantitative information. It is worth noting that all SSP scenarios result in an emissions pathway higher than RCP 4.5, a radiative forcing level that was previously considered ‘middle of the road’ (IPCC, 2014). Three out of five scenarios are described in Table 6 below. The most sustainable future is described by SSP1 and is characterized by a swift transition to clean energy, low resource demand, and globally cooperative nations that enable adaptation and mitigation of climate change (O’Neill et al., 2017; Riahi et al., 2017). Under this scenario, the worst impacts of climate change are avoided, and global cooperation supports efforts to improve health and wellbeing, ultimately implying reduced risk to NTM and other opportunistic infections. SSP2 and SSP3 provide less optimistic futures for the environment and for NTM, mainly due to an inability to manage finite water resources and innovate or distribute management solutions equitably.

**TABLE 6. SSPs considered with respect to NTM.** Descriptions for the SSPs were obtained from O’Neill et al. (2017) and Riahi et al. (2017) and represent a few of many potential pathways and outcomes. Interpretations for NTM are an extension of the literature review and discussion in the previous sections.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Description</th>
<th>Potential implications for NTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSP1</td>
<td>Cooperative governance, sustainable development, urbanization, equity, healthcare and education access, lower population, clean energy, innovation for environmental technologies, sharp cut in resource demand. Low mitigation and adaptation challenges. Radiative forcing ~5 W/m²</td>
<td>Cooperation reduces de facto reuse by embracing a One Water approach, reducing pathogen exposure. Innovation and equity bring improved water treatment and reclamation practices to more people, reducing NTM exposure. Temperatures continue to rise, supporting NTM in most natural environments, except for those with increased salinity. Heightened focus on wellbeing and equity reduces vulnerability and adverse infection outcomes during disaster times.</td>
</tr>
<tr>
<td>SSP2</td>
<td>No major changes from status quo, slow movements towards sustainable development,</td>
<td>Lack of major innovation and persistent inequality suggest that improved water treatment and management strategies fail to</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
innovation proceeds slowly without major breakthroughs, resource demand reduces slowly, inequality persists, urbanization remains at pace and does not accelerate.

### Medium challenges to mitigation and adaptation

Radiative forcing ~ 6.5 W/m²

reach all populations, allowing de facto reuse and existing inadequate treatment options to persist in practice. Pressure on water resources continues, resulting in greater reliance on reclaimed water, but without new management practices, potentially exposing more people to NTM. Continuing urbanization trends allow greater impervious cover and increase urban heat island effect, exacerbating eutrophication and flood challenges in highly populated areas, potentially exposing more people to NTM in affected water sources and flood water.

### SSP3

Regional rivalries, resurgent nationalism, focus on national security at the expense of environmental issues, inequality persists, technological investments decline, population grows in developing countries and stabilizes in industrial, urbanization and economic growth stabilize.

High challenges to mitigation and adaptation

Radiative forcing ~ 7.2 W/m²

Climate change accelerates with a high level of radiative forcing, exacerbating warming, humidity, evapotranspiration, sea level rise, natural disaster intensity and/or frequency, extreme precipitation and drought, and all of the associated effects on water scarcity and NTM. Nationalism prevents cooperative water management, stressing water resources and increasing reliance on inadequately treated reclaimed water. Inequalities prevent increases in education and access to healthcare, increasing vulnerability, and resulting in worse outcomes for those exposed to NTM.

3.5. Uncertainties and recommendations for future work

As with any projections and scenario analyses, there are many uncertainties involved in the scenarios described and in their interpretation with respect to NTM. Uncertainties involved in modeling due to randomness and gaps in knowledge impact climate change projections as well as predictions of NTM abundance under different conditions.

Uncertainties also exist with respect to the choices that humans will make as well as the ultimate effect of those choices. For example, in SSP1, innovation is key and there are uncertainties as to how successful innovative efforts may be at tackling challenges including energy efficiency and water quality/quantity. There are uncertainties about how social behaviors and interactions with the environment (e.g., time spent outside) will change in the future, which may also influence vulnerability and relative risk. Although robust modeling efforts by IPCC scientists attempt to delineate the extent of uncertainties, there will always be...
potential surprises given the large number of interacting factors involved in all climate and SSP projections.

More data collection in a variety of geographic locations spanning climates and socioeconomic statuses is needed. For example, study is warranted of potential relationships between thawing permafrost over peatlands in boreal regions (Turetsky et al., 2000) and NTM concentrations in connected freshwater systems. Previous reports have associated NTM abundance with peatland drainage areas (Iivanainen et al., 1993) and the organic acids found in peat (Kirschner et al., 1999), but no conclusive evidence exists regarding how permafrost thaw might affect that dynamic or human populations downstream. More broadly, current knowledge about climate change and water quality is based on few studies that consider limited variables, and focus on surface water in developed countries (IPCC, 2014). A gap exists in understanding how much groundwater hydrology and quality has already been affected by climate change, land use, and water abstraction patterns due to a lack of adequate monitoring wells and studies (IPCC, 2014).

With respect to NTM fate and transport, there are also uncertainties about the interactions taking place between environmental variables and matrices. These uncertainties can be better understood with more study of NTM in connected water and air environments over extended time periods and diverse environmental conditions. Further research is needed to fill the remaining gaps about the relationships between NTM, soils, temperature, nutrients, metals, salinity, and oxygen, and how some of these factors exert influences in engineered systems. Additionally, more research is recommended about the relationships between evapotranspiration and NTM transport through air to understand why evapotranspiration is related to disease risk in some regions (Adjemian et al., 2012). Understanding the mechanisms behind NTM survival and transport in aerosols in arid climates where rates of evapotranspiration are projected to increase would inform the potential for changes in risk.
Future research could also consider what effect changes in UV irradiation (Bais et al., 2015) may have on NTM, given the known differences in UV susceptibility across species (David, 1973). Additional research investigating the safety of reclaimed water and aquifer recharge for pathogens including NTM is also needed, and studies focusing on relative risk of these water conservation practices as compared to no action would be particularly useful for decision making. More research is also needed about potential reclaimed water treatment and regrowth control options. Some topics that are recommended for future study include the potential for UV to disinfect without causing concerning mutations, use of alternative corrosion inhibitors to protect pipes and prevent biofilms while not adding nutrients, and options for pipe protection (e.g., shade and minimum clearance distances) to control temperatures within distribution systems. Future work should also investigate the practicality of establishing advanced treatment processes across the globe and determine what innovative solutions are needed.

Models that seek to identify the critical limits of factors that influence NTM fate and transport, and quantify the uncertainties in natural and engineered systems would be a natural next step towards understanding changing risk. Modeling expected impacts to viable NTM cells rather than gene concentrations could enhance utility. Research that seeks to understand how gene concentrations translate to viable cells may help in that effort.

More research is also needed to better understand exposure routes and how they may be impacted by climate change. This requires more work involving aerosolization efficiency and transport. Collaborations between patients, medical professionals, environmental microbiologists, engineers, and climatologists may facilitate a more complete understanding of likely transmission routes, exposure levels, and dose-response parameters necessary for QMRA. A public health reporting requirement for NTM infections does not exist in most of
the world, and if implemented, would enhance the understanding of these relationships by providing more data on the prevalence of NTM infections.

4. Conclusions

Climate change may lead to a net increase in NTM exposure from water sources due to their adaptability to environmental conditions imposed by warming temperatures as well as their ability to be easily aerosolized. Infection risk may also increase due to their opportunistic nature that may benefit from an increasingly stressed population. Engineering solutions exist, such as reverse osmosis and nanofiltration, that could reduce exposure to NTM from reclaimed water sources, but more research is needed to understand regrowth during distribution of reclaimed water and potential solutions, as well as the practicality of making these solutions accessible worldwide. Equitable and sustainable socioeconomic choices may provide an opportunity for water managers to collaboratively manage water resources through One Water practices and innovation that reduce contamination by unmanaged wastewater, reducing risk of infection by pathogens including NTM. More field data is needed to understand impacts of climate change on water quality broadly, and to specifically understand NTM fate and transport as a function of environmental factors, as well as relevant exposure routes. Collaborative studies that merge knowledge across fields of expertise would support better understanding of relative risk of infection by pathogens including NTM in a complex global change landscape.
IV. CONCLUSION

As NTM infections and their associated costs continue trending upwards, it is increasingly urgent that their preferred habitats at the human-environment interface be better understood. Little is known about NTM prevalence in private wells and the home biofilm microbiome in homes using private well water, which nearly 43 million people rely on in the US. In the research presented here, NTM were detected for the first time in New Jersey (NJ) plumbing biofilms in homes using private wells, and disinfection selection pressure did not significantly affect their detection. Using random forest (RF) regression, physiography and within-home location, corresponding to different use patterns and environments, were the most important variables that explained variations in mycobacteria gene relative abundance. Using linear discriminant effect size analysis, one group of nonpathogenic biomarkers was associated with the absence of mycobacteria genes, indicating a potential direction for future studies of ecological engineering. It is recommended that future studies obtain a greater sample size that may facilitate more robust analyses. Future studies should also consider viability- or activity-based analyses to understand how much of the observed abundance is viable. Consideration of environmental factors not included in this study, such as water age, may increase the percent of variation that can be explained by models such as RF. Private well homes should be included in future studies of NTM in home plumbing biofilms as they serve millions of people and can contain similar abundances of NTM as publicly supplied homes.

As environmental bacteria often present in water environments, NTM are expected be affected by climate, engineering, and societal changes. Although researchers have studied NTM dynamics and environmental relationships, and have written reviews that summarize their ecology and epidemiology, no critical reviews have been compiled that consider the multitude of inter- and intra-system interactions that may determine prevalence and exposure under different future scenarios. In evaluating climate, engineering, and societal forces, this review found that
NTM in many water environments may find more suitable habitats with warmer temperatures and associated water quality changes (e.g., eutrophication) owing to their adaptability. However, sea level rise and associated increases in salinity upstream may reduce habitat suitability for NTM at the saltwater/freshwater interface. Engineering solutions exist, such as reverse osmosis and nanofiltration, that could reduce exposure to NTM from reclaimed water sources, but more research is needed to understand regrowth during distribution of reclaimed water. Equitable distribution of advanced treatment systems and equitable water management practices that embrace One Water solutions to reduce or mitigate the impacts of de facto reuse may be important control measures as water scarcity requires more reliance on reclaimed water. Without intervention, infection risk may also increase due to the opportunistic nature of NTM that may find enhanced prospects for infection in an increasingly stressed population under climate change. More field data is needed to understand impacts of climate change on water quality, and to specifically understand NTM fate and transport as a function of multiple environmental factors. Collaborative studies that merge knowledge across fields of expertise would support better understanding of the relative risk of infection by pathogens including NTM in a complex global change landscape.
V. APPENDIX: Ch II Supplementary Materials

S1 DNA extraction method modification details

Prior to extracting DNA from study samples, the extraction method was tested on biofilm swabs taken from sinks within our lab on public water supply. While processing these extracts for downstream analyses, the extracts were observed to have a gelatinous texture, potentially indicating polysaccharide co-precipitation (Fang et al., 1992). To remedy this issue, the lysis buffer sodium chloride concentration was increased from 0.24 M to 1 M and a second ethanol wash step was included (Fang et al., 1992). Following the modification, all extracts maintained their aqueous state.

S2 LefSe LDA analysis continued

For private well biofilms, two biomarker families were Gemmatimonadaceae and Rhodocyclaceae, families of Gram-negative bacteria with differing characteristics and environmental prevalence between them. Gemmatimonadaceae have been found in many environments including agricultural soil, hypoxic cave sediment, freshwater, UV treated drinking water and wastewater (Martin-Pozas et al., 2020; Pullerits et al., 2020; Rodriguez-Sanchez et al., 2018; Soinne et al., 2020; Zhao et al., 2020). Likewise, members of the beta-proteobacterial Rhodocyclaceae family have been identified in many environments including sediments, contaminated soils, freshwater lakes, and aquifers (Corteselli et al., 2017; Táncsics et al., 2018). The Rhodocyclaceae family includes species with a variety of phenotypic properties, and the genera identified in this study were majority (84%) denitrifying bacteria, Denitratisoma, and 15% nitrifying bacteria, Methyloversatilis, though these should be taken with caution as they are based on short reads. The Phycisphaeraceae family was also identified as a biomarker for private wells, and is a family that has been found in wastewater, and that has exhibited antibiotic resistance (Wang et al., 2020).
**TABLE S1** qPCR primers, reaction conditions, reaction metrics, and amplicon lengths. Values are reported as average ± standard deviation. LOD and LOQ are log_{10}(x+1) copies per swab.

<table>
<thead>
<tr>
<th>Target Gene</th>
<th>Primer sequence (5’–3’</th>
<th>Thermocycler conditions</th>
<th>R²</th>
<th>E</th>
<th>Matrix Spike Log % Recovery</th>
<th>LOD</th>
<th>LOQ</th>
<th>Amplicon length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total bacteria 16S rRNA (Muyzer et al., 1993)</td>
<td>CCTACGGGAGGCAGCAG ATTACCGCGGCTGCTGG</td>
<td>95 °C for 10m [95 °C for 15s, 66.4 °C for 6s] x³⁹</td>
<td>0.99</td>
<td>5 ± 0.00</td>
<td>98 ± 10%</td>
<td>100.2 ± 2.2%</td>
<td>5.50</td>
<td>5.50</td>
</tr>
<tr>
<td>M. avium 16S rRNA (Wang et al., 2012)</td>
<td>AGAGTTTGATCCTGGCTCA G ACCAGAAGACATGCGTCT TG</td>
<td>98 °C for 2m [98 °C for 5s, 68 °C for 18s] x⁴₀</td>
<td>0.99</td>
<td>6 ± 0.00</td>
<td>85 ± 5%</td>
<td>98.3 ± 3.1%</td>
<td>-</td>
<td>4.44</td>
</tr>
<tr>
<td>atpE (Radomski et al., 2013)</td>
<td>CGGYGCCGGTATCGGYGA CGAAGACGAACARSGCCA T FAM/ACSGTGATGAAGACCATGCTGCTCACAGTTA AAACCTTGTTACCGCGGCTGCTGG</td>
<td>95 °C for 45s [95 °C for 3s, 64.4 °C for 30s, 60 °C for 30s] x⁴₀</td>
<td>0.99</td>
<td>5 ± 0.00</td>
<td>85 ± 7%</td>
<td>96.9 ± 2.3%</td>
<td>4.02</td>
<td>4.67</td>
</tr>
<tr>
<td>Mycobacterium spp. 16S rRNA (myco16S) (Radomski et al., 2010)</td>
<td>CCTGGGAACCTGGGTCTA AT CGCACGCTCACAGTTA FAM-TTCACGAACACGCGAC AAACCTTGTTACCGCGGCTGCTGG</td>
<td>95 °C for 2m [95 °C for 5s, 55 °C for 15s, 72 °C for 10s] x⁴₅</td>
<td>0.99</td>
<td>7 ± 0.00</td>
<td>90 ± 2%</td>
<td>100.7 ± 1.6%</td>
<td>2.69</td>
<td>4.53</td>
</tr>
</tbody>
</table>

*E = Efficiency*
**TABLE S2** Aggregated survey responses and sampling details. All survey and sampling data analyzed is included here. Samples from each group included in the sequencing analysis are also detailed. Samples sent for sequencing analysis were chosen to best represent the variation in the samples collected, with a particular focus on physiography and pipe material.

<table>
<thead>
<tr>
<th>Parameter recorded</th>
<th># of biofilm samples collected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(# analyzed by amplicon sequencing)</td>
</tr>
<tr>
<td><strong>Physiographic region</strong></td>
<td></td>
</tr>
<tr>
<td>Coastal plain:</td>
<td>15 (5)</td>
</tr>
<tr>
<td>Piedmont:</td>
<td>42 (17)</td>
</tr>
<tr>
<td>Highlands:</td>
<td>10 (5)</td>
</tr>
<tr>
<td>Valley &amp; Ridge:</td>
<td>3 (2)</td>
</tr>
<tr>
<td><strong>Pipe material</strong></td>
<td></td>
</tr>
<tr>
<td>Copper:</td>
<td>39 (12)</td>
</tr>
<tr>
<td>PVC:</td>
<td>6 (3)</td>
</tr>
<tr>
<td>Copper &amp; PVC:</td>
<td>19 (11)</td>
</tr>
<tr>
<td>PVC &amp; PEX:</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Unknown:</td>
<td>3 (1)</td>
</tr>
<tr>
<td><strong>Water supply</strong></td>
<td></td>
</tr>
<tr>
<td>Private well:</td>
<td>59 (21)</td>
</tr>
<tr>
<td>Public supply:</td>
<td>11 (8)</td>
</tr>
<tr>
<td><strong>Sampling location</strong></td>
<td></td>
</tr>
<tr>
<td>Showerhead:</td>
<td>23 (0)</td>
</tr>
<tr>
<td>Sink:</td>
<td>47 (29)</td>
</tr>
<tr>
<td><strong>Disinfection method</strong></td>
<td></td>
</tr>
<tr>
<td>Chlorine:</td>
<td>14 (9)</td>
</tr>
<tr>
<td>UV:</td>
<td>9 (1)</td>
</tr>
<tr>
<td>No disinfection:</td>
<td>44 (18)</td>
</tr>
<tr>
<td>Unknown:</td>
<td>3 (1)</td>
</tr>
<tr>
<td><strong>Filter used?</strong></td>
<td></td>
</tr>
<tr>
<td>Yes:</td>
<td>32 (10)</td>
</tr>
<tr>
<td>No:</td>
<td>35 (18)</td>
</tr>
<tr>
<td>Unknown:</td>
<td>3 (1)</td>
</tr>
<tr>
<td>Survey question</td>
<td>Yes</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td><strong>Water softener installed?</strong></td>
<td>33</td>
</tr>
<tr>
<td>Yes- softener:</td>
<td></td>
</tr>
<tr>
<td>No- softener:</td>
<td></td>
</tr>
<tr>
<td>Unknown:</td>
<td></td>
</tr>
<tr>
<td><strong>Well age</strong></td>
<td></td>
</tr>
<tr>
<td>&lt; 50 years:</td>
<td>27</td>
</tr>
<tr>
<td>≥ 50 years:</td>
<td>23</td>
</tr>
<tr>
<td>Unknown:</td>
<td>9</td>
</tr>
<tr>
<td><strong>Wellhead visible aboveground?</strong></td>
<td></td>
</tr>
<tr>
<td>Yes:</td>
<td>47</td>
</tr>
<tr>
<td>No:</td>
<td>12</td>
</tr>
<tr>
<td><strong>Soil raised around the wellhead?</strong></td>
<td></td>
</tr>
<tr>
<td>Yes:</td>
<td>12</td>
</tr>
<tr>
<td>No:</td>
<td>41</td>
</tr>
<tr>
<td>Unknown:</td>
<td>6</td>
</tr>
<tr>
<td><strong>Septic tank?</strong></td>
<td></td>
</tr>
<tr>
<td>Yes:</td>
<td>16</td>
</tr>
<tr>
<td>No:</td>
<td>4</td>
</tr>
<tr>
<td><strong>Sampling time</strong></td>
<td></td>
</tr>
<tr>
<td>Morning:</td>
<td>49</td>
</tr>
<tr>
<td>Afternoon:</td>
<td>9</td>
</tr>
<tr>
<td>Evening:</td>
<td>12</td>
</tr>
<tr>
<td><strong>Sampling season</strong></td>
<td></td>
</tr>
<tr>
<td>Spring:</td>
<td>9</td>
</tr>
<tr>
<td>Summer:</td>
<td>39</td>
</tr>
<tr>
<td>Fall:</td>
<td>17</td>
</tr>
<tr>
<td>Winter:</td>
<td>5</td>
</tr>
<tr>
<td><strong>Sampling notes</strong></td>
<td></td>
</tr>
<tr>
<td>Easy sampling:</td>
<td>46</td>
</tr>
<tr>
<td>Swabbed exterior:</td>
<td>16</td>
</tr>
<tr>
<td>No aerator:</td>
<td>3</td>
</tr>
<tr>
<td>Aerator up-pipe:</td>
<td>3</td>
</tr>
<tr>
<td>Heavy rust:</td>
<td>2</td>
</tr>
</tbody>
</table>

*a* Survey questions & sampling details

*b* Samples that maintained $\geq 10^4$ sequences after processing through QIIME

*c* “Easy sampling” defined as when all attachments could be removed

*d* “Swabbed exterior” defined as when the aerator or showerhead could not be removed
TABLE S3 Average 16S rRNA gene amplicon sequences per sample returned and filtered prior to analysis. Samples with ≤ 10^2 sequences were excluded from the rarefied analyses due to low sequence counts. Due to losses during filtering, one sample with original sequence counts ≥ 10^4 is not represented in the table or in the 22,000 sequences subsampling depth. This sample fell between the two major groups below, with 10^3 filtered sequences. Notably, showerhead biofilm samples returned uniformly low sequence counts, with the maximum being 86 total unfiltered sequences.

<table>
<thead>
<tr>
<th>Sample group</th>
<th>Average total sequences/sample</th>
<th>Average remaining filtered &amp; denoised sequences/sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 10^2 sequences a (N = 31)</td>
<td>52 ± 57</td>
<td>5 ± 21</td>
</tr>
<tr>
<td>≥ 10^4 sequences a (N = 29)</td>
<td>100918 ± 55727</td>
<td>56438 ± 26805</td>
</tr>
</tbody>
</table>

a Final, processed sequence counts
### TABLE S4 NTM qPCR and sequencing results by household.*

<table>
<thead>
<tr>
<th>Home ID</th>
<th>Location</th>
<th>atp / 16S rRNA gene copies</th>
<th>myco16S / 16S rRNA gene copies</th>
<th>Mycobacteriaceae sequences / total</th>
<th>atpE +/-</th>
<th>myco16S +/-</th>
</tr>
</thead>
<tbody>
<tr>
<td>WG1</td>
<td>Kitchen B2</td>
<td>-2.71</td>
<td>-6.10</td>
<td>-2.53</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.41</td>
<td>-3.05</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>WG2</td>
<td>Alt. Sink B2</td>
<td>-5.25</td>
<td>-4.89</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Bathroom 2</td>
<td>-5.91</td>
<td>-3.55</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-1.84</td>
<td>-3.66</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>WG3</td>
<td>Alt. sink B2</td>
<td>-2.86</td>
<td>-3.80</td>
<td>-2.59</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-5.16</td>
<td>-6.10</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WG4</td>
<td>Kitchen</td>
<td>-5.51</td>
<td>-5.15</td>
<td>-2.42</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-4.14</td>
<td>-5.08</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WG5</td>
<td>B1</td>
<td>-1.80</td>
<td>-1.53</td>
<td>-0.57</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-2.23</td>
<td>-3.72</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-1.22</td>
<td>-1.79</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>WG6</td>
<td>Kitchen B2</td>
<td>-3.94</td>
<td>-4.88</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.90</td>
<td>-3.54</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>WG7</td>
<td>B1</td>
<td>-4.63</td>
<td>-5.57</td>
<td>-2.66</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-2.14</td>
<td>-2.16</td>
<td>-1.21</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-4.80</td>
<td>-5.74</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WG8</td>
<td>B1</td>
<td>~ -2.71</td>
<td>~ -4.14</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-2.29</td>
<td>~ -4.52</td>
<td>-2.48</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-4.66</td>
<td>-5.60</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>WG9</td>
<td>B1</td>
<td>-5.94</td>
<td>-1.59</td>
<td>-0.33</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-5.34</td>
<td>-2.39</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.91</td>
<td>~ -4.01</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>WG10</td>
<td>Kitchen B2</td>
<td>-5.05</td>
<td>~ -4.69</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-2.08</td>
<td>-4.98</td>
<td>-</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>WG11</td>
<td>B1</td>
<td>-3.84</td>
<td>-1.59</td>
<td>-0.30</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-5.77</td>
<td>-1.72</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>~ -4.70</td>
<td>~ -4.80</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>WG12</td>
<td>Kitchen B2</td>
<td>-5.56</td>
<td>~ -5.20</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-1.60</td>
<td>-1.83</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>B2</td>
<td>Showerhead</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------------</td>
<td>-------------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WG13</td>
<td>-4.01</td>
<td>-4.95</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5.44</td>
<td>-6.38</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.75</td>
<td>~ -4.39</td>
<td>-</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>-5.16</td>
<td>-</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-5.16</td>
<td>-2.59</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-1.36</td>
<td>-1.05</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>-1.32</td>
<td>-1.34</td>
<td>-0.37</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-4.20</td>
<td>-5.14</td>
<td>-2.58</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.17</td>
<td>-4.11</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Alt. sink</td>
<td>-2.97</td>
<td>-3.91</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-2.81</td>
<td>-3.75</td>
<td>-2.77</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.88</td>
<td>-0.60</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Kitchen a</td>
<td>~ -2.85</td>
<td>~ -3.79</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-2.54</td>
<td>-3.48</td>
<td>-2.66</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead a</td>
<td>~ -2.61</td>
<td>~ -3.55</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>-1.05</td>
<td>-1.55</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-5.53</td>
<td>~ -5.17</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.00</td>
<td>-3.03</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>-3.06</td>
<td>-4.00</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-3.49</td>
<td>-4.43</td>
<td>-2.69</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.14</td>
<td>-4.54</td>
<td>-</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>-4.87</td>
<td>~ -4.51</td>
<td>-2.79</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-3.66</td>
<td>-4.60</td>
<td>-2.64</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.67</td>
<td>~ -3.31</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>-4.46</td>
<td>~ -4.10</td>
<td>-2.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-2.80</td>
<td>~ -2.44</td>
<td>-0.69</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-4.93</td>
<td>-2.43</td>
<td>-</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>-3.62</td>
<td>~ -3.26</td>
<td>-2.46</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B2</td>
<td>-2.76</td>
<td>-3.70</td>
<td>-2.65</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-3.28</td>
<td>-4.22</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>-5.76</td>
<td>-3.37</td>
<td>-2.37</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>-5.89</td>
<td>-3.78</td>
<td>-2.35</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Showerhead</td>
<td>-2.60</td>
<td>-2.22</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Kitchen a</td>
<td>~ -2.56</td>
<td>~ -3.50</td>
<td>-2.65</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>B1</td>
<td>-2.85</td>
<td>-3.79</td>
<td>-2.46</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Samples marked with * had 16S rRNA gene copies < LOQ
Patient home on private well water
Patient home on public water sourced from groundwater
* All data is log10(x+1) transformed. “~” identifies gene copy values below LOD or LOQ, which were substituted with LOD/√2 or LOQ/√2, respectively. B1 = first floor bathroom sink, B2 = second floor bathroom sink, kitchen = kitchen sink, Alt. sink = basement or other sink.
TABLE S5 Measures of alpha diversity of biofilm sample amplicon sequences, subsampled to 22,000 sequences per sample, separated by water supply. Values are reported as average ± standard deviation.

<table>
<thead>
<tr>
<th>Alpha diversity index</th>
<th>Private well ((n = 21))</th>
<th>Public Supply ((n = 8))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richness</td>
<td>66 ± 25</td>
<td>49 ± 17</td>
</tr>
<tr>
<td>Shannon (H')</td>
<td>2.32 ± 0.46</td>
<td>1.98 ± 0.30</td>
</tr>
<tr>
<td>Evenness ([H' / \ln (\text{richness})])</td>
<td>0.56 ± 0.11</td>
<td>0.51 ± 0.16</td>
</tr>
</tbody>
</table>
**TABLE S6** Results of linear discriminant analysis (LDA) effect size (LefSe) using default parameters.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>Biomarker bacterial family</th>
<th>Higher classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>myco16S</td>
<td>Absent</td>
<td><em>Family XI</em></td>
<td>Bacilli</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Rikenellaceae</em></td>
<td>Bacteroidetes</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Pseudomonadaceae</em></td>
<td>Gammaproteobacteria</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Erysipelotrichaceae</em></td>
<td>Firmicutes</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td><em>Beijerinckiaeeae</em></td>
<td>Alphaproteobacteria</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Spiromaceae</em></td>
<td>Bacteroidetes</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Chitinophagaceae</em></td>
<td>Bacteroidetes</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Isosphaeraceae</em></td>
<td>Planctomycetes</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Acetobacteraceae</em></td>
<td>Alphaproteobacteria</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Reyranellaceae</em></td>
<td>Alphaproteobacteria</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Hyphomicrobiaceae</em></td>
<td>Alphaproteobacteria</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Methylpilaceae</em></td>
<td>Alphaproteobacteria</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Rhizobiaceae</em></td>
<td>Alphaproteobacteria</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Gemmatimonadaceae</em></td>
<td>Gemmatimonadetes</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Rhodocyclaceae</em></td>
<td>Betaproteobacteria</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Phyctisphaeraceae</em></td>
<td>Planctomycetes</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Xanthobacteraceae</em></td>
<td>Alphaproteobacteria</td>
</tr>
<tr>
<td></td>
<td>Private well</td>
<td><em>Thermaceae</em></td>
<td>Deinococcus–Thermus</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Enterobacteriaceae</em></td>
<td>Gammaproteobacteria</td>
</tr>
<tr>
<td></td>
<td>Public supply</td>
<td><em>Muribaculaceae</em> (formerly S24-7)</td>
<td>Bacteroidetes</td>
</tr>
</tbody>
</table>
FIG S1 Alpha rarefaction curves (N = 22,000 seqs/sample) split into first floor bathroom and kitchen sinks (a) and second floor bathroom sinks (b) for ease of viewing. The nine homes represented with two samples match in color for both charts and can be identified by the last 1–2 digits. Sample names coded as W = private well, G = general public, P = public supply, N = patient; b1 = first floor bathroom sink, b2 = second floor bathroom sink, k = kitchen sink, a = alternative sink, remaining 1–2 digits = system ID (1–20 for private wells; 1–4 for public supply).
FIG S2 qPCR marker gene log abundance (a) \( atpE \) and (b) myco16S by physiography. Samples with abundances < LOD or < LOQ are substituted with LOD/√2 or LOQ/√2, respectively.
FIG S3 Random Forest measures of variable importance for (a) 16S rRNA gene copy, (b) myco16S gene relative abundance, (c) Shannon’s $H'$, (d) species richness, and (e) evenness regression models. Starting with 24 total, variables were iteratively removed if they negatively influenced the accuracy of the model, as indicated by a negative value for “% increase in mean squared error.” The original variable set included the following: (1) physiographic region (2) pipe material (3) private well v. public supply (4) sample location (5) disinfection system (6) filter usage (7) water softener usage (8) salt added to the softener (9) sampling time (10) sampling season (11) sampling notes (12) well age (13) wellhead aboveground (14) raised soil around wellhead (15) ground v. surface water (different than public v. private due to one public supply that uses groundwater) (16) maintenance history (17) septic tank usage (18) presence of permanent structures 10’ from wellhead (19) pH (20) conductivity (21) temperature (22) residual chlorine concentration (23) total coliform presence (24) E. coli presence. Co-correlated variables with a Spearman’s rho coefficient > 0.5 were not included in the same model.
VI. REFERENCES


Blanc, S., Pender, D., Vinnard, C., Gennaro, M.L. and Fahrenfeld, N.L. In revision Nontuberculous mycobacteria in the biofilm microbiome of private well and premise plumbing.


Breiman, L. 2002 Manual on setting up, using, and understanding random forests v3.1.


Cui, Q., Fang, T., Huang, Y., Dong, P. and Wang, H. 2017. Evaluation of bacterial pathogen diversity, abundance and health risks in urban recreational water by amplicon next-generation sequencing and quantitative PCR. Journal of Environmental Sciences 57, 137-149.


Donohue, M.J. 2018. Increasing nontuberculous mycobacteria reporting rates and species diversity identified in clinical laboratory reports. BMC Infectious Diseases 18(1), 163.


García-Coca, M., Rodríguez-Sevilla, G., Pérez-Domingo, A., Aguilera-Correa, J.-J., Esteban, J. and Muñoz-Egea, M.-C. 2020. Inhibition of *Mycobacterium*
abscessus, M. chelonae, and M. fortuitum biofilms by Methylobacterium sp. The Journal of Antibiotics 73(1), 40-47.


Luan, J., Zhang, C., Xu, B., Xue, Y. and Ren, Y. 2020. The predictive performances of random forest models with limited sample size and different species traits. Fisheries Research 227, 105534.


Parker, B.C., Ford, M.A., Gruft, H. and Falkingham III, J.O. 1983. Epidemiology of infection by nontuberculous mycobacteria: IV. Preferential aerosolization of


Pogoda, K.N.E. (2017) Targeted 16S rRNA metagenomic analysis of E. coli contaminated private drinking well waters in southern Ontario, Queen's University, Kingston, Ontario, Canada.


Are there environmental or agricultural benefits in using forest residue biochar in boreal agricultural clay soil? Science of the Total Environment 731, 138955.


Thomson, R.M., Carter, R., Tolson, C., Coulter, C., Huygens, F. and Hargreaves, M. 2013b. Factors associated with the isolation of Nontuberculous mycobacteria (NTM) from a large municipal water system in Brisbane, Australia. BMC microbiology 13(1), 89.


USEPA 2020 National Primary Drinking Water Regulations.


VII. ACKNOWLEDGEMENT OF PREVIOUS PUBLICATIONS

Chapter II is in review for publication. The citation for this pending publication is below. The author contributions are as follows:

Conception/ design of the work: CV, NLF, MLG. Data collection: SB, DP. Data analysis & interpretation: SB, NLF. Drafting the article: SB. Critical revision of the article: SB, CV, MLG, NLF.

Blanc, S., Pender, D., Vinnard, C., Gennaro, M.L., and Fahrenfeld, N.L. Nontuberculous mycobacteria in the biofilm microbiome of private well and premise plumbing. *In revision.*