Cast iron drinking water pipe biofilms support diverse microbial communities containing antibiotic resistance genes, metal resistance genes, and class 1 integrons†

Lee K. Kimbell, Emily Lou LaMartina, Anthony D. Kappell, Jingwan Huo, Yin Wang, Ryan J. Newton and Patrick J. McNamara

Antimicrobial resistance is a well-documented public health concern. The role that drinking water distribution pipes have as sources of antibiotic resistance genes (ARGs) is not well known. Metals are a known stressor for antibiotic resistance development, implying that aging metal-pipe infrastructure could be a source of ARGs. The objective of this study was to determine if ARGs, metal resistance genes (MRGs), and \textit{intI1} were pervasive across various pipe biofilm sample types (biomass surfaces, pipe surfaces, corrosion tubercles, and under corrosion tubercles) and if the resistance genes associated with particular microbial taxa. Eight sample types in triplicate \((n=24)\) were taken from inside a \(>100\) year-old, six ft. section of a full-scale chloraminated cast iron drinking water main. Droplet digital PCR (ddPCR) was employed as a novel approach to quantify ARGs in pipes from full-scale drinking water distribution systems (DWDS) because it yielded higher detection frequencies than quantitative PCR (qPCR). Illumina sequencing was employed to characterize the microbial community based on 16S rRNA genes. ARGs and MRGs were detected in all 24 pipe samples. Every sample contained targeted genes. Interestingly, the mean absolute abundances of ARGs and MRGs only varied by approximately one log value across sample types, but the mean relative abundances (copy numbers normalized to 16S rRNA genes) varied by over two log values. The ARG and MRGs concentrations were not significantly different between sample types, despite significant changes in dominant microbial taxa. The most abundant genera observed in the biofilm communities were \textit{Mycobacterium} (0.2–70%), and \(\beta\)-lactam resistance genes \textit{bla}\textsubscript{TEM}, \textit{bla}\textsubscript{SHV}, and the integrase gene of class 1 integrons (\textit{intI1}) were positively correlated with \textit{Mycobacterium}. The detection of ARGs, MRGs, and class 1 integrons across all sample types within the pipe indicates that pipes themselves can serve as sources for ARGs in DWDS. Consequently, future work should investigate the role of pipe materials as well as corrosion inhibitors to determine how engineering decisions can mitigate ARGs in drinking water that stem from pipe materials.

1. Introduction

Antibiotic resistance is a major public health concern stemming from the microbial response to the widespread occurrence of antibiotics and other physiological stressors in the environment.\(^1\)\(^2\) Approximately 2.8 million people are diagnosed with infections caused by antibiotic-resistant bacteria (ARB), and over 35,000 deaths are attributed to antibiotic resistance annually in the U.S. alone.\(^3\) Antibiotic
resistance genes (ARGs) on mobile genetic elements can be acquired by pathogens in the human gut, creating risks for vulnerable populations that are exposed to ARGs. ARGs have been detected in various water environments including groundwater, surface water, drinking water treatment plants, and tap drinking water at concentrations up to $10^{10}$ copies per L. Consequently, it is becoming increasingly important to quantify ARGs in exposure routes that directly convey ARGs to people, including drinking water distribution systems (DWDS).

ARGs in tap water have been shown to increase from the drinking water treatment plant effluent to the tap. While residual disinfectants can select for antibiotic resistance, the actual infrastructure of DWDS, i.e., the pipe materials used, could also be an important factor that impacts microbial ecology and consequently ARG profiles. DWDS are comprised of a variety of metallic pipe materials (e.g., copper, iron, and lead) and additional metals in treated drinking water can accumulate in biofilms and corrosion scales. Metals select for antibiotic resistance through co-resistance and cross-resistance mechanisms. Additionally, microorganisms have evolved detoxification strategies, such as metal resistance genes (MRGs) and efflux pumps, to mitigate the toxic effects of metals. Exposing bacteria to metals in DWDS may promote the survival of bacteria resistant to metals and antibiotics. While studies have documented the occurrence of ARGs in tap drinking water and in biofilms, to the best of our knowledge, no research efforts have quantified ARGs, MRGs, and mobile genetic elements from different biofilm sample locations (e.g., surface biofilms, tubercles, under tubercles) in a single pipe to understand if drinking water pipes can serve as sources of ARGs.

The objective of this research was to determine if ARGs, MRGs, and class 1 integrons (intI1) were quantifiable across multiple sample types in a chloraminated cast iron water main and to determine if microbial taxa were correlated to resistance gene concentrations. It was hypothesized that ARGs, MRGs, and intI1 would be detected regardless of sample type and location. The abundance of bacterial biomass (measured by 16S rRNA gene copies), ARGs, MRGs, and the integrase gene of the class 1 integron, intI1, were quantified in samples collected from different microenvironments using droplet digital PCR (ddPCR) and quantitative PCR (qPCR). Microbial communities were analyzed using PCR-amplified 16S rRNA gene sequences from each pipe sample. This is the first research to determine if various types of biofilm samples from a single full-scale DWDS pipe can serve as potential sources for ARGs.

2. Materials and methods

2.1 Pipe collection, sampling, and DNA extraction

A six ft section of cast-iron water main (18” ID, 105 years in operation) that transported chloraminated water was extracted, covered with sterile plastic sheeting, and immediately transported to the laboratory for sampling and analysis. The water main was collected as part of planned maintenance to replace old water mains throughout the distribution system. Pipe samples were collected from 1) a visible biofilm surface, referred to as “biomass surface” (n = 6), ii) a pipe surface that did not have biofilm visible to the naked eye, referred to as “pipe surface” (n = 6), iii) from three-dimensional corrosion tubercles that could be removed, referred to as “corrosion tubercles” (n = 6), and iv) from the pipe surface on the location where the tubercle was removed, referred to as “under corrosion tubercles” (n = 6) (ESI† Fig. 1). Broadly speaking, all samples were microbial biofilms, not bulk water samples, and they were subcategorized into the four categories listed for comparison and statistical analysis. Pipe surface, biomass surface, and under tubercle swabs were collected by firmly pressing a sterile cotton-tipped applicator (Fisher Scientific, Waltham, MA) on the biofilm surface and swabbing an area of approximately 2–5 cm$^2$. For each of the microenvironment types sampled, top (n = 3) and bottom (n = 3) samples of the cast iron water main were collected. Each swab was transferred directly to a sterile DNA extraction lysing tube and the stem was snapped and severed to preserve only the sample end of the swab. Tubercle samples were collected into plastic tubes using a flame sterilized spatula, and approximately 0.2 g of corrosion tubercle was sub-sampled for DNA extraction. Samples were immediately frozen at −20 °C until DNA extraction was performed. DNA was extracted using the FastDNA Spin Kit (MP Biomedicals, Solon, OH). The manufacturer’s protocol was followed with the exception

![Fig. 1 X-ray diffraction (XRD) patterns of biofilm tubercle samples (n = 6) collected from the chloraminated cast-iron water main. The reference XRD patterns for goethite (JCPDS 29-0713) and lepidocrocite (JCPDS 08-0098) are also shown for comparison.](image-url)
that initial cell lysis was conducted using liquid nitrogen freeze thaw cycling (3×). DNA concentrations in resulting extracts were quantified by microspectrophotometry using a Nano-Drop (Nano-Drop™ Lite, Thermo Scientific, Waltham, MA) and stored at −20 °C.

### 2.2 Quantification of resistance genes

Droplet digital PCR (ddPCR) assays were conducted to quantify gene copies. A subset of samples was initially analyzed at 5, 10, 50, and 100-fold dilutions to test for inhibition during gene quantification. Based on these results, a 10-fold dilution was selected to minimize inhibition for all samples. Reaction mixtures consisted of a total volume of 22 µL with 11 µL of QX200 ddPCR EvaGreen Supermix (final concentration 1×) (Bio-Rad Laboratories Inc., Hercules, CA), 2 µL each of forward and reverse primers (final concentration 250 nM each), 4 µL of diluted DNA extract, and 3 µL molecular grade water. The ddPCR reaction mixture was added to a 96-well plate, sealed with foil, homogenized by vortexing, and centrifuged briefly to ensure that all reaction components were at the bottom of the wells. The 96-well plate was equilibrated at room temperature for 3 minutes prior to droplet generation. Aliquots of 20 µL for each reaction were dispensed into a separate well of an eight-channel droplet generator cartridge (DG8 Cartridge, Bio-Rad) followed by 70 µL of QX200 Droplet Generation Oil for EvaGreen into the oil wells for subsequent droplet generation using the QX100 Droplet Generator. Oil-droplet mixtures were transferred to a 96-well plate and sealed at 180 °C using the PX1 PCR plate sealer. The 96-well plate was transferred to the C1000 touch thermal cycler for PCR thermal cycling with the following conditions: 5 min at 95 °C for activation of DNA polymerase, 39 cycles of 95 °C for 30 s and 60 °C for 60 s, followed by signal stabilization at 4 °C for 5 min and 90 °C for 5 min. Thermal cycling conditions were modified for genes with annealing temperatures varying from 60 °C (ESI† Table S1). After thermal cycling, plates were transferred to the QX200 Droplet Reader for absolute quantification of target genes.

Data analysis was performed using the QuantaSoft Analysis Pro software and expressed as gene copies µL⁻¹ (V 1.0.596, Bio-Rad). Positive controls were included with each ddPCR assay and were produced by ten-fold serial dilution of plasmid DNA yielding 10⁴ to 10⁹ copies per reaction. No-template (i.e., reagent only) controls were included with each ddPCR assay. All ddPCR negative controls failed to yield amplification above the limit of quantification for each assay. Thresholds to discriminate between positive and negative droplets were manually applied to each sample and only samples with ≥3 positive droplets were considered as positive. Furthermore, only reactions with greater than 10,000 accepted droplets were used for subsequent analysis.

The limit of the detection (LOD) and limit of quantification (LOQ) for each tested gene were determined according to the MIQE guidelines. Additional MIQE guidelines were followed and are shown in ESI† Table S2. DNA extract from each sample was analyzed in triplicate for each target gene, and the average value from analyzing each DNA sample three times was used for each sample for further analysis.

Target gene copies were also quantified in triplicate from DNA extracts using qPCR with previously published protocols for the 16S rRNA gene, ARGs (bla<sub>TEM</sub>, bla<sub>SHV</sub>, sul<sub>1</sub>, cop, SHV,10 bla<sub>TEM</sub>, sul<sub>1</sub>,16 MRGs (czcD, copA), and the integrase gene of class 1 integrons (intI1). Additional information on qPCR methodology, specific primer sets, amplicon sizes, annealing temperatures, R² values, efficiencies, and quantification limits are described in the ESI† Methods and Table S1. β-lactamase resistance genes such as bla<sub>TEM</sub> and bla<sub>SHV</sub> are grouped in the most common types of β-lactamases belonging to Enterobacteriaceae and encode resistance to β-lactam antibiotics such as penicillins and cephalosporins. Sulphonamide resistance gene (sul1) and the integrase gene of class 1 integrons (intI1) are frequently detected in various natural and engineered environments and are considered a good proxy for ARG abundance and anthropogenic pollution.50,51 MRGs quantified in biofilm samples from the cast iron water main included the copper resistance gene copA and czeD, which is a part of the cation diffusion facilitator mediating resistance to cadmium, zinc, and cobalt. These genes were selected based on their abundance in an initial qPCR assay conducted with over 20 different MRGs including genes encoding resistance to metals such as arsenic, copper, iron, lead, and zinc.

### 2.3 PCR and Illumina sequencing of 16S rRNA gene amplicons

Microbial communities from biofilm samples were prepared for analysis by triplicate PCR-amplifying and pooling V4 hypervariable regions of 16S rRNA genes. One extraction blank and mock community (#HM-782D, BEI Resources) were included in the sample set. PCR amplicons were sequenced with Illumina MiSeq 2 × 250 paired-end chemistry at the Great Lakes Genomic Center (http://greatlakesgenomics.uwm.edu). Primer and barcode sequences were removed from raw reads using cutadapt. Reads were processed, including filtered, merged, error-corrected, and chimera-checked, into amplicon sequence variants (ASVs) using the R package phyloseq. ASVs that were classified as mitochondria, chloroplast, or eukaryota were removed. Additional thresholds were set to identify and remove ASVs potentially derived from the mock community, extraction/PCR blank, and non-target samples that were included in the sequencing run. Additional description of sequencing methods is included in the ESI† Methods. Raw sequences have been uploaded to NCBI Sequence Read Archive (SRA) under BioProject ID PRJNA692495.

### 2.4 Corrosion tubercle characterization

X-ray diffraction (XRD) analysis was conducted on corrosion tubercles that were sampled from the cast iron water main (n = 6) to identify the dominant crystalline phases. XRD was performed on a Bruker D8 Discover A25 diffractometer using
copper Kα radiation with step scanning from 2θ of 15–70°. The scan speed and step size were 3° per min and 0.02°, respectively. XRD patterns of each corrosion tubercle were compared to standard patterns from International Centre for Diffraction Data (ICDD).

The inorganic elemental composition of corrosion tubercles was determined using inductively coupled plasma mass spectrometry (ICP-MS).58 Approximately 0.1 g of each corrosion tubercle was subsampled for elemental analysis. Each tubercle sample was homogenized using a sterile mortar and pestle prior to acid digestion with nitric acid (2%) and hydrochloric acid (1%).59 An Agilent Technologies 7700 Series ICP-MS (Agilent Technologies Inc., Santa Clara, CA) was used for elemental composition determination. Standard reference materials for elements including Ag, Al, As, Ba, Be, Ca, Cd, Co, Cr, Cu, Fe, Hg, Mg, Mn, Mo, Na, Ni, Pb, Sb, Se, Th, Ti, U, V, and Zn were purchased from Agilent Technologies.

### 2.5 Statistical analysis

Statistical analyses were conducted using RStudio in the open-sourced statistical program R (V 3.6.1).60,61 One-way analysis of variance (ANOVA) was conducted using the ‘aov’ function to determine statistical differences between abundances of target genes across groups of samples. A significant cutoff of α = 0.05 was used for all analyses. For sequence analysis, the BIOM file generated via DADA2 was imported into R using the phyloseq package.62 R packages ‘phyloseq’ and ‘ggplot2’ were utilized for general visualization of sequence data. Alpha and beta diversity metrics and plots were generated using the ‘vegan’ and ‘ggplot2’ packages. ANOVA was used to determine significance among the alpha diversity metrics. Principal coordinate analysis (PCoA) was performed using the ‘ape’ package to visualize differences between samples using the Bray–Curtis dissimilarity matrix generated in ‘phyloseq’. Canonical correspondence analysis (CCA) was conducted in R using the vegan package to identify correlations between the bacterial community structure and biofilm sample location. Spearman’s rank sum correlation coefficients were calculated in R to assess correlations between ARGs, MARGs, abundant taxa, and biofilm sample location. Indicator taxa were identified for each sample location using the multi-level pattern (indicator species) analysis in the package ‘indicspecies’.63 Rarefaction curves were generated using the ‘ggregr’ function from the phyloseq-extended package of scripts.64

### 3. Results and discussion

#### 3.1 Physical characterization of corrosion tubercles

XRD analysis indicated that the mineral phases present in the interior surfaces of the cast iron pipe primarily consisted of goethite (α-FeOOH) and lepidocrocite (γ-FeOOH) (Fig. 1). Goethite has been widely found as a main corrosion product in cast iron pipes65–70 and several previous studies also reported the observation of lepidocrocite in iron pipes from full-scale DWDS.71,72 The inorganic elemental composition of corrosion tubercules was characterized by ICP-MS. Iron was the dominant element in corrosion tubercules representing approximately 98.6% of the measured mass. Other elements detected in quantities ranging from 0.1% to 1.0% in the tubercules included Ca (0.64%), Al (0.38%), Mg (0.12%), and Mn (0.10%). Elements detected below 0.1% included Na, K, Co, Cu, Pb, Zn, V, As, Se, Mo, Ag, Cd, Hg, Ni, and Be. Metals present in corrosion scales and tubercules in the cast iron water main may have originated from the pipe material itself, especially Fe, but other trace elements were likely deposited over time from the bulk drinking water.

#### 3.2 Quantification of bacterial biomass

16S rRNA genes were detected above quantification limits in all 24 samples from the cast iron water main (ESI† Fig. S2). The mean concentration in corrosion tubercles was $4.4 \times 10^7$ 16S rRNA gene copies per g tubercle. The mean concentration in biofilms collected under corrosion tubercles was $3.5 \times 10^5$ 16S rRNA copies per cm$^2$. Previous studies have reported similar levels of 16S rRNA in biofilms from chloraminated water mains with averages ranging from $3.2 \times 10^5$ to $2.5 \times 10^7$ copies per cm$^2$.73–75

#### 3.3 Quantification of resistance genes in pipe samples

##### 3.3.1 Detection frequency by ddPCR and qPCR

The frequency of gene detects for every gene analyzed across the 24 samples using ddPCR was equal to or higher than that for qPCR (Table 1; ESI† Fig. S3). The presence of inhibiting substances such as metals or humic acids in the biofilm samples are known to impact amplification and primer annealing in qPCR assays.76,77 Previous studies have also demonstrated that ddPCR, as compared to qPCR, can have increased precision and accuracy for quantifying low concentrations of DNA in variably contaminated samples.76–78 Our findings in conjunction with previous studies suggest that ddPCR is favorable for detecting ARGs in DWDS, particularly because these samples often contain low concentrations of DNA and contaminants that may interfere with qPCR. Reporting limit of detection and limit of quantification will be key for comparing across studies, as these values can differ significantly among quantification methods.

##### 3.3.2 Abundance of antibiotic resistance genes (ARGs) and intI1

The ARGs $bla_{ARM}$, $bla_{TEM}$, and $sul_1$, and intI1 were detected in biofilm samples from the chloraminated cast-iron drinking water main at concentrations up to nearly 6 log gene copies cm$^{-2}$, with the highest mean value belonging to gene $bla_{TEM}$ at approximately 4.8 log gene copies cm$^{-2}$ (Fig. 2). A previous study reported the mean absolute abundance of ARGs ($bla_{TEM}$, $sul_1$, $qnr_3$, $vanA$) and intI1 in biofilms from undefined pipe materials to range from <LOQ
Mean absolute abundances of ARGs and \(\text{intI}1\) in different biofilm microenvironments varied by over one log unit, but the mean differences of the sample types were not significantly different from each other (one-way ANOVAs, \(p\) values > 0.05), indicating that sample type did not impact absolute gene abundances. Relative gene abundances (absolute normalized to 16S rRNA gene copies) demonstrated higher variability (>2-log units) between the different microenvironments sampled, but the mean relative abundance values were not significantly different based on Table 1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Mechanism/mode of action</th>
<th>ddPCR detections</th>
<th>ddPCR LOQ (CN (\mu L^{-1}))</th>
<th>ddPCR LOQ (CN cm(^{-2}))</th>
<th>qPCR detections</th>
<th>qPCR LOQ (CN (\mu L^{-1}))</th>
<th>qPCR LOQ (CN cm(^{-2}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA</td>
<td>NA</td>
<td>24 (100%)</td>
<td>5 (\times 10^3)</td>
<td>24 (100%)</td>
<td>500</td>
<td>5 (\times 10^3)</td>
<td></td>
</tr>
<tr>
<td>(\text{bla}_{\text{SHV}})</td>
<td>Beta lactam resistance</td>
<td>24 (100%)</td>
<td>2 (\times 10^4)</td>
<td>20 (83%)</td>
<td>50</td>
<td>5 (\times 10^3)</td>
<td></td>
</tr>
<tr>
<td>(\text{bla}_{\text{TEM}})</td>
<td>Beta lactam resistance</td>
<td>24 (100%)</td>
<td>2 (\times 10^4)</td>
<td>24 (100%)</td>
<td>5</td>
<td>5 (\times 10^3)</td>
<td></td>
</tr>
<tr>
<td>copA</td>
<td>Copper resistance</td>
<td>11 (45%)</td>
<td>4 (\times 10^4)</td>
<td>3 (12.5%)</td>
<td>5</td>
<td>5 (\times 10^3)</td>
<td></td>
</tr>
<tr>
<td>(\text{czcD})</td>
<td>Cobalt, zinc, cadmium efflux</td>
<td>20 (83%)</td>
<td>3 (\times 10^4)</td>
<td>19 (79%)</td>
<td>5</td>
<td>5 (\times 10^3)</td>
<td></td>
</tr>
<tr>
<td>(\text{intI}1)</td>
<td>Integrase gene of class 1 integrons</td>
<td>10 (42%)</td>
<td>4 (\times 10^4)</td>
<td>8 (33%)</td>
<td>5</td>
<td>5 (\times 10^3)</td>
<td></td>
</tr>
<tr>
<td>sul1</td>
<td>Enzymatic modification</td>
<td>17 (71%)</td>
<td>3 (\times 10^4)</td>
<td>4 (17%)</td>
<td>5</td>
<td>5 (\times 10^3)</td>
<td></td>
</tr>
<tr>
<td>(\text{tet}(L))</td>
<td>Tetracycline efflux</td>
<td>0 (0%)</td>
<td>6 (\times 10^4)</td>
<td>0 (0%)</td>
<td>50</td>
<td>5 (\times 10^4)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: CN – copy numbers, LOQ – limit of quantification, NA – not applicable.

Fig. 2 Absolute abundance of antibiotic resistance genes, metal resistance genes, and \(\text{intI}1\) in different biofilm microenvironments from a cast iron drinking water main as measured with ddPCR. The biofilm microenvironments include biomass surface (BS), pipe surface (PS), tubercle (TUB), and under tubercle (UT). Each biofilm sample is also categorized by top or bottom pipe sample location. Sample categories (e.g. bottom UT, top BS) are plotted when 1–3 of the samples yielded a quantifiable result. Sample categories with no positive detections were left blank. Each symbol represents the average of 3 ddPCR technical replicates from a single biofilm sample. All biofilm swab samples are plotted as \(\log_{10}\) gene copies per cm\(^2\). Corrosion tubercle samples are plotted as \(\log_{10}\) gene copies per g. The quantification limit (QL) is also plotted for each gene as the dash on the right side of the plot.
sample type (one-way ANOVA, \( p > 0.05 \)) (ESI† Fig. S4). The variability in relative abundance values for ARGs in the different microenvironments was primarily due to differences in levels of bacterial 16S rRNA genes between sample locations rather than changes in ARG abundance. Indeed, these results indicate that various pipe biofilm samples, and thus pipe infrastructure, could serve as sources of ARGs into tap drinking water.

Absolute concentrations of ARGs remained relatively consistent in the different biofilm sample locations. ARGs were detected more frequently in surficial biofilm environments but demonstrated similar absolute abundance compared to sub-surface environments. One explanation could be that corrosion deposits and tubercles can provide relief from disinfectants, advective flow, and shear stress which allows additional biofilm development to occur.\(^{40} \)

Observed relative abundance values for ARGs in the current study suggest that microbes in sub-surface communities may harbor ARGs at similar levels compared to surficial microbes. The presence of ARGs in each of the different biofilm sample locations suggests that potential selection for antibiotic resistance exerted by disinfectants, metals and other dissolved contaminants exists throughout the cast iron biofilm communities inside the pipe. Additionally, the detection of ARGs and class 1 integrons in the biofilm communities suggests that horizontal gene transfer may be one plausible explanation for the proliferation of ARGs inside the cast iron biofilm communities due to the close proximity of microbial species within biofilms.\(^{30,79} \)

Average and median chloramine concentrations inside the distribution system were 1.0 to 1.3 mg L\(^{-1} \). Residual disinfectants such as chloramines present in DWDS at subinhibitory levels have dissolved contaminants exists throughout the cast iron biofilm communities due to the close proximity of microbial species within biofilms.\(^{30,79} \)

Average and median chloramine concentrations inside the distribution system were 1.0 to 1.3 mg L\(^{-1} \). Residual disinfectants such as chloramines present in DWDS at subinhibitory levels have been shown to stimulate horizontal gene transfer of ARGs through multiple pathways including reactive oxygen species response systems and the SOS response.\(^{18} \)

Beta-lactamase resistance genes, \( \text{bla}_{\text{SHV}} \) and \( \text{bla}_{\text{TEM}} \), were detected in all 24 biofilm samples ranging from 4.1 to 5.3 log gene copies per cm\(^2 \). A previous study detected \( \text{bla}_{\text{TEM}} \) in drinking water biofilms at lower absolute abundance (mean = 1.5 log gene copies swab\(^{-1} \)) but with similar relative abundance (mean = -2.3 log ARG copies/16S rRNA copies) as in this study.\(^{34} \)

These results demonstrate that different microenvironments or niches in pipes can serve as reservoirs for bacteria harboring resistance genes, indicating that biofilms on DWDS pipes can serve as sources of ARGs when biofilms shed viable cells into tap drinking water.\(^{79} \)

The sulfonamide resistance gene (\( \text{sul1} \)) and the integrase gene of class 1 integrons (\( \text{int1} \)) are frequently detected in various natural and engineered environments and are considered potential indicators of horizontal gene transfer of ARGs.\(^{50,51} \)

The \( \text{int1} \) gene was detected in 10 biofilm samples at an average concentration of 4.3 log gene copies per cm\(^2 \). The \( \text{int1} \) gene was detected at a higher frequency in biofilm surface samples (83%) and was not detected in tubercle samples. The frequency of \( \text{sul1} \) detections for biofilm surface, pipe surface, and under tubercle samples was 83%, compared to 33% for corrosion tubercles. Additional variation in ARG abundance was observed between samples collected from the top and bottom of the water main, but the differences were not statistically different (one-way ANOVA, \( p \) values >0.05). The distance of biofilm sample collection inside the water main (12 in., 18 in., and 24 in.) also did not have a significant effect on ARG concentrations.

### 3.3.3 Abundance of metal resistance genes (MRGs)

MRGs quantified in biofilm samples from the cast iron water main included the copper resistance gene \( \text{copA} \) and \( \text{czcD} \). The \( \text{czcD} \) gene was detected in 83% of biofilm samples at concentrations ranging from above LOQ to 4.6 log gene copies cm\(^{-2} \). The mean \( \text{czcD} \) absolute abundance in surficial samples (BS, PS) was not significantly different compared to sub-surface samples (TUB, UT) according to One-way ANOVA results (\( p > 0.05 \)). The \( \text{czcD} \) gene is part of the cation diffusion facilitator mediating resistance to cadmium, zinc, and cobalt\(^{52,53} \) and has previously been documented in source drinking water treated with the corrosion inhibitor zinc orthophosphate.\(^{58} \)

The copper resistance gene \( \text{copA} \) was detected above the LOQ in 45% of samples with ddPCR assays. The \( \text{copA} \) gene encodes an ATPase efflux pump that extrudes copper ions, making it one of the main mechanisms mediating copper resistance.\(^{83} \)

The detection of \( \text{czcD} \) and \( \text{copA} \) may be related to the presence of copper, zinc, cadmium, and cobalt ions in the drinking water and corrosion deposits. Previous studies have demonstrated that exposure to sub-inhibitory levels of \( \text{Cu(n)} \) and \( \text{Zn(n)} \) can contribute to horizontal gene transfer of ARGs.\(^{29} \)

This is the first study to quantify clinically-relevant ARGs and MRGs in multiple types of biofilm samples from a chloraminated cast iron water main collected from a full-scale distribution system.

### 3.4 Microbial community composition of pipe samples

Corrosion tubercles, especially from the bottom of the water main, exhibited the most unique microbial community composition compared to the biomass surface, pipe surface, and under tubercle samples (Fig. 3a). Biofilm microbial communities exhibited similar Shannon diversity (1.8–4.4) to previous observations of biofilm communities in cast iron drinking water mains\(^{73} \) (ESI† Fig. S5). A total of 469 microbial genera corresponding to 47 different phyla were identified in the biofilm microbial communities from the cast iron water main. The most abundant genera observed in the biofilm communities included \( \text{Myccobacterium} \) (0.2–70%), \( \text{Geobacter} \) (0–57%), \( \text{Gallionella} \) (0–40%), \( \text{Phreatobacter} \) (0–25%), \( \text{Desulfovibrio} \) (0–21%), \( \text{Undibacterium} \) (0–18%), \( \text{Streptococcus} \) (0–17%), and \( \text{Sphingomonas} \) (0–17%) (Fig. 3b). Previous studies have also observed high abundance of \( \text{Mycobacterium} \), \( \text{Geobacter} \), \( \text{Gallionella} \), \( \text{Sphingomonas} \), and \( \text{Undibacterium} \) in corrosion deposits and biofilms in DWDS.\(^{58,82–84} \)

Other abundant genera observed in the cast iron biofilm communities included \( \text{Hydrogenophaga} \) (0–15%), \( \text{Rhodofex} \) (0–15%), \( \text{Galbitalea} \) (0–12%), \( \text{Corynebacterium} \) (0–11%), \( \text{Ralstonia} \) (0–8%), and \( \text{Geothrix} \) (0–6%).
Mycobacterium represented the most abundant genus in the current study with 22 unique Mycobacterium ASVs detected in the biofilm communities. These ASVs comprised 32% of the total sequences in the biofilm samples and were detected in all 24 biofilm samples. Mean relative abundance of Mycobacterium spp. was highest in under tubercle samples (43%) compared to other sample locations (25–32%), but the differences were not statistically significant (one-way ANOVA, p > 0.05). Although the mean values were not significantly different, relative abundance of Mycobacterium exhibited wide variability across individual samples ranging from 0.2 to 70%. These findings are consistent with previous studies that have reported Mycobacterium as the dominant genus in drinking water biofilms from chloraminated cast iron water mains. Mycobacterium spp. are frequently detected in DWDS and have several characteristics that give them a competitive advantage in chloraminated DWDS including the ability to form biofilms, resistance to residual disinfectants, and the ability to survive in nutrient-deficient environments. The prevalence of Mycobacterium in full-scale DWDS is a potential concern because the genus contains several opportunistic pathogens.
The abundance patterns for the most abundant taxa are shown in Fig. 4. Several corrosion-related bacterial genera including *Sphingomonas*, *Desulfovibrio*, *Gallionella*, *Geobacter*, *Hydrogenophaga*, and *Rhodobacter* were observed.
in the biofilm communities and have been previously
documented in cast-iron pipe biofilms and corrosion
scales.67,68,73,94–96 Bacteria inhabiting distribution systems
primarily perform functions related to carbon source utilization
and nitrification, as well as microbial induced corrosion
processes.97 Corrosion tubercles contained increased
abundance of microorganisms related to microbial induced
corrosion and redox cycling processes in cast iron pipes
including genera such as Gallionella, Geobacter, Geothrix, and
Undibacterium.67,68,73 Previous studies have demonstrated that
the abundance of microorganisms in DWDS may be impacted
by the presence of corrosion tubercles, suspended solids, and
other loose corrosion deposits.98,99 Additionally, the bottom of
water mains can accumulate higher densities of bacteria and
corrosion deposits, which may also support increased
abundance of nitrifying bacteria.100 Similarly, the biofilm
samples from the bottom of the cast iron water main displayed
increased relative abundance of bacterial genera related to iron
and nutrient cycling.

Spearman correlation analysis revealed significant
relationships between the co-occurrence of several corrosion-
related bacteria in the biofilm communities (ESI† Fig. S6). The
genus Ralstonia demonstrated the most significant
relationships with other taxa, and it is considered an
emerging global opportunistic pathogens in municipal water
supplies.101 Positive relationships were observed between the
co-occurrence of corrosion-related bacteria from Geobacter
and Gallionella genera, which also made up a large
proportion of the communities in corrosion tubercles. Geobacter were most prevalent in corrosion tubercles (11.7%)
but were also identified in lower abundance in pipe surface
(3.1%), biofilm surface (0.4%), and under tubercle
environments (0.4%). The Geobacter genus contains iron-
reducing species, and previous studies have reported that
Geobacter spp. were among the most resistant bacteria to
monochloramine disinfection in lab-scale disinfection
experiments.102 Similarly, Gallionella spp. were more
abundant in corrosion tubercles (16.5%) compared to biofilm
surface (3.6%), pipe surface (1.0%), and under tubercle
environments (1.9%). Gallionella spp. are neutrophilic iron
oxidizing bacteria (IOB) that can promote the precipitation of
iron oxides by converting ferrous iron to ferric iron and have
previously been observed in association with severe iron
corrosion release or “red water” events in distribution
systems.96,103 Bottom tubercle samples demonstrated
increased relative abundance of Gallionella and Geobacter
(22.5% and 23%) compared to top tubercle samples (10.5%
and 0.4%), but these differences were not significant (one-
way ANOVA, p values >0.05).

Indicator species analysis was employed to identify the
most abundant indicator organisms in each sample location
from the water main. Several corrosion-related genera were
significant indicator organisms for corrosion tubercle
samples including Geothrix, Gallionella, Phreatobacter,
Thiimonas, and Rhodovastum (all p values <0.05). Methylobacterium were identified as significant indicators for
biofilm surface communities, while Bradyrhizobium were
identified as indicators of pipe surface and under tubercle
communities. Phreatobacter and Ralstonia genera were
identified as indicators of communities in samples collected

![Fig. 5] Relationships between relative abundance of most abundant taxa and absolute abundance of ARGs, MRGs, and intI1 in biofilm samples from cast iron water main. Color denotes the result from correlation analysis using Spearman's rank sum correlation with Spearman's rho value plotted for each comparison. The rho value for statistically significant relationships are also included (p values <0.05). The lowest available taxonomic classification for each observed ASV is provided.
from pipe surfaces, biofilm surfaces, and under tubercles. Several of the identified indicator organisms in our study have been reported in previous studies analyzing cast iron biofilm communities.\(^{68,83,104}\)

### 3.5 Relationships between sample location, resistance genes, and microbial community composition

Microbial communities and abundances of resistance genes were compared between sample locations to determine potential host bacteria harboring ARGs, MRGs, and int\text{}\textsuperscript{I}.\(^{105}\) Additionally, the relationships among bacterial community composition and different sample types were evaluated using CCA. Biofilm sample type significantly influenced microbial community composition and explained 32\% of the community variability (CCA, \(p = 0.009\)). Biofilm samples were taken from 12, 18, and 24 in. inside of the water main to assess impact of lateral sampling distance. CCA indicated that sampling distance into the pipe explained 36\% of the variability in bacterial community composition (\(p = 0.001\)). These results indicate that biofilm sample location was a significant variable influencing microbial communities.

Spearman correlation analysis was conducted to determine relationships between bacterial community composition and abundance of resistance genes. Significant correlations were observed between several microbial genera and the abundance of resistance genes in the biofilm samples (Fig. 5). The occurrence of \(czc\)D was significantly correlated with genera including Desulfovibrio, Ferrireducticum, Herbaspirillum, and Rhodoferax (all \(p \text{ values} < 0.05\)). Positive relationships were also observed between \(czc\)D and corrosion-related genera such as Geothrix, Gallionella, Sphingomonas, and Undibacterium.

The copper resistance gene \(copA\) was positively associated with the occurrence of several genera including Gallitalea, Gallionella, Geobacter, Geothrix, and Ferrireducticum. A significant negative relationship was observed between \(copA\) and Phreatobacter. \(\beta\)-Lactam resistance genes \(bla_{\text{TEM}}\) and \(bla_{\text{SHV}}\) were positively correlated with taxa including Burkholderia, Gallitalea, and Mycobacterium. A significant negative relationship was observed between \(bla_{\text{TEM}}\) and the genus Hydrogenophaga (\(p < 0.05\)). The integrase gene of class 1 integrons (\(intI\)) was positively correlated with genera including Desulfovibrio, Gallitalea, Hydrogenophaga, and Mycobacterium. The sulfonamide resistance gene \(sul1\) demonstrated a significant positive relationship with the Geobacter genus.\(^{106}\) The presence of statistically significant correlations between microbial taxa and resistance genes in the biofilm samples implies that at least some of the shifts observed in gene abundance could have resulted from shifts in abundances or types of host bacteria.\(^{34}\)

### 4. Conclusions

This research established that drinking water main biofilms in a chloraminated cast iron water main can serve as sources of resistance genes, regardless of sample type or location within the pipe. This is the first research to use ddPCR and qPCR to quantify ARGs, MRGs, and class 1 integrons in drinking water biofilms from a full-scale distribution system. Future research should be conducted to determine the distribution and concentrations of resistance genes in different pipe materials and in different locations of the same distribution system. Additional research is needed to quantify resistance genes and microbial communities in other pipe materials including copper, lead, and plastic that are commonly used in full-scale DWDS to understand how engineering management decisions can impact sources of antibiotic resistance.\(^{34}\) Further, full-scale and laboratory scale studies should be conducted to determine the impacts of corrosion inhibitors and corrosion products on the prevalence of antibiotic resistance in DWDS. Previous studies have suggested that drinking water biofilms may facilitate horizontal transfer of the ARGs from one host to another due to the presence of nutrients and high bacterial density and diversity.\(^{8}\) In the current study, genera containing opportunistic pathogens detected in the biofilm samples included Mycobacterium, Ralstonia, Staphylococcus, and Sphingomonas. Further research targeting these specific bacterial genera would be necessary to determine the presence of any potential opportunistic pathogens. Human exposure routes relevant for potable water include consumption of tap drinking water, dermal contact, and inhalation of aerosolized drinking water during showering or bathing.\(^{16}\) Studies documenting the occurrence and distribution of ARGs and MRGs in DWDS are critical for human health risk assessments evaluating the potential for the transfer and development of antibiotic resistance in engineered systems and in the environment.\(^{105,107}\) Given the potential for bacterial growth in DWDS, it is essential to continue to seek water treatment and management options that minimize levels of antibiotic resistance. Specific conclusions from key findings are as follows:

- ARGs (\(bla_{\text{TEM}}, \text{ bla}_{\text{SHV}}, \text{ sul1}\)), MRGs (\(copA, \text{ czcD}\)), and/or class 1 integrons (\(intI\)) were detected in every biofilm sample type studied within a chloraminated cast iron drinking water main, indicating that pipes could serve as sources for ARGs.
- ddPCR assays resulted in more positive detections and lower detection limits for target genes compared to qPCR assays. Future studies should consider ddPCR for environmental samples containing inhibiting substances such as metals, humic acids, and other contaminants.
- Microbial communities varied between different biofilm sample locations and were dominated by corrosion-related genera including Mycobacterium, Geobacter, Gallionella, and Sphingomonas.
- Significant relationships were observed for the co-occurrence of ARGs, MRGs, \(intI\) and several microbial taxa.
- Further research is warranted to determine the impacts of different pipe materials on the abundance of ARGs, MRGs, and \(intI\) in biofilms inhabiting full-scale DWDS.

### Conflicts of interest

There are no conflicts to declare.
Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant Numbers 2027288 and 2027233. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We would like to thank Dr. Walter McDonald for his assistance with transporting the pipe to the lab. We also thank Mike Dollhopf and the Marquette University (MU) Water Quality Center for analysis of metals using the ICP-MS. L. K. K. was supported by an Arthur J. Schmitt Fellowship from Marquette University. E. L. M. was supported by a student research fellowship from the Milwaukee Metropolitan Sewerage District. This research was supported by a Regular Research Grant from the MU Office of Innovation for ddPCR equipment and supplies. The Way Klinger Young Scholar Sabbatical Award supported P. J. M. and sequencing costs. We acknowledge the use of XRD at the Advanced Analysis Facility at the University of Wisconsin-Milwaukee. The authors acknowledge the use of the LightCycler 96 for qPCR from Marquette University, funded by the GHR Foundation.

References


Environmental Science: Water Research & Technology

Environmental Science: Water Research & Technology


45 E. Marti, J. Jofre and J. L. Balcazar, Prevalence of Antibiotic Resistance Genes and Bacterial Community Composition in a River Influenced by a Wastewater Treatment Plant, PLoS One, 2013, 8(10), e78906, DOI: 10.1371/journal.pone.0078906.


M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBnet.journal, 2011, 17(1), 10–12.


Environmental Science: Water Research & Technology


87 H. Sun, B. Shi, Y. Bai and D. Wang, Bacterial community of biofilms developed under different water supply conditions in a distribution system, Sci. Total Environ., 2014, 472, 99–107, DOI: 10.1016/j.scitotenv.2013.11.017.


