METAGENOMIC ANALYSIS OF ANTIBIOTIC RESISTANT GENES IN A CONVENTIONAL AND MEMBRANE BIOREACTOR WASTEWATER TREATMENT PLANT

By

Camille McCall

A THESIS

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

Environmental Engineering – Master of Science

2016

ABSTRACT

METAGENOMIC ANALYSIS OF ANTIBIOTIC RESISTANT GENES IN A CONVENTIONAL AND MEMBRANE BIOREACTOR WASTEWATER TREATMENT PLANT

By

Camille McCall

Wastewater treatment plants (WWTPs) are known environments for the presence and transfer of antibiotic resistant genes (ARGs), an evolving environmental pollutant. This study aimed to explore the prevalence of ARGs and resistant bacteria in a conventional, and a membrane bioreactor (MBR) WWTP in Michigan (USA). A sequence-based metagenomic approach was implemented to detect the profile of ARGs in the activated sludge (AS), before disinfection (BD), and after disinfection (AD) treatment stages in each WWTP. Metagenomic alignment detected genes resistant to sulfonamide, tetracycline, macrolide, elfamycin, aminoglycoside, and β-lactam classes of antibiotics to be prevalent ARGs in both WWTPs. Effluent samples yielded the highest presence of ARGs in each plant compared to AS and BD samples. Quantitative analysis found that 56.25% and 53.33% of a total of 23 ARGs, which were detected at ≥ 90% gene similarity among all samples, were not detected until after disinfection samples for the conventional and MBR WWTPs, respectively. To our knowledge, this study is the first to report the prevalence of elfamycin resistant genes in WWTPs. In addition to this, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Pasteurella multocida* were found to be predominant resistant bacteria in AD samples from each WWTP. The occurrence of ARGs increased in both WWTPs as treatment progressed further suggesting that increased wastewater treatment selects for antimicrobial resistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The development and overuse of antibiotics has led to the selection of antibiotic resistant genes (ARGs) in bacteria resulting in reduced susceptibility to a wide range of antimicrobial treatments. Antibiotics are largely used in animal production and are among the largest therapeutic treatment for bacterial infections in humans (Manaia et al., 2015; Schmieder and Edwards, 2012). It is said that within the first few years of introducing a new antibiotic, pathogens commonly known to persist in hospital settings, develop resistance (Schmieder and Edwards, 2012; Zhang et al., 2015b). To date, nosocomial (hospital) infections are one of the leading causes of death (Xia et al., 2016). Resistant pathogens of this nature escape controlled settings through, most commonly, hospital waste streams and are released into the environment. Antibiotic resistant genes have become an evolving environmental pollutant known to occur in ecosystems such as, soils, surface waters, and WWTPs (Manaia et al., 2015). WWTPs are known to select for antimicrobial resistance given the various treatment processes and the harboring of an abundance of diverse microbial communities (Mao et al., 2015; Murray et al., 1984).

While ARGs in WWTPs have been explored in many facets, the effects of treatment on these genes are still unclear. Several studies report discrepancies in the effects of treatment on the behavior of ARGs throughout various wastewater treatment stages (Diehl and LaPara, 2010; Murray et al., 1984). The differences in how ARGs behave throughout waste treatment cycles can be caused by several factors that are at play. For example, biological treatment types, hydraulic detention time, presence of heavy metals, and temperature all play a role in the prevalence and removal of ARGs in wastewater treatment systems (Bondarczuk et al., 2016; Diehl and LaPara, 2010; Gao et al., 2012b; Mao et al., 2015; Novo and Manaia, 2010; Zhang et al., 2015c; Zhang et al., 2015b). Common ARGs found in waste streams such as tetracycline and sulfonamide resistant

genes have been studied at great lengths, generally by means of culturing and polymerase chain reaction (PCR) techniques (Mao et al., 2015, Gao et al., 2012a; Zhang et al., 2011; Gao et al., 2012b; Diehl and LaPara, 2010; Novo and Manaia, 2010; Munir et al., 2011). However, studies capturing a wide spectrum of ARGs in various types of wastewater treatment systems are limited. Considering the role of WWTPs in the selection or reduction of ARGs, the need to investigate the prevalence and diversity of these genes on a metagenomic scale has become increasingly important over the years.

Though, culture-based techniques have made it possible to detect ARGs in environmental samples, a reported 99% of environmental bacteria cannot be cultured (Varela and Manaia, 2013). This makes culture-base metagenomics insufficient in identifying novel, or a broad spectrum of antibiotic resistant bacteria within microbial communities (Schmieder and Edwards, 2012). The introduction of sequence-based metagenomics allows for the detection of a wide range of ARGs through genomic mapping and gene affiliation to national reference databases (Schmieder and Edwards, 2012). Sequence-based techniques involve sequencing a random sample of genetic material recovered directly from the environment without the use of culturing (Penders et al., 2013). Sequencing platforms such as Illumina, produce vast libraries of genetic information that are later sequenced and can then be post-processed for further analysis.

In this study, we used a sequenced-based metagenomic approach to assess the occurrence of ARGs in a conventional activated sludge (CAS) with chlorine (Cl) disinfection, and a membrane bioreactor (MBR) with ultraviolet (UV) disinfection WWTP in Michigan. Samples were extracted from the activated sludge, before disinfection, and after disinfection stages of each WWTP. Our

major goal is to identify predominant antibiotic resistant bacteria (ARB) in AD samples from each WWTP.

2. MATERIALS AND METHODS

2.1. Sample Collection

Sewage samples were collected from the East Lansing and Traverse City WWTPs in Michigan (U.S.A.) in 2013. The characteristics of these WWTPs are shown in Table 2.1. Samples were taken from three different locations throughout the treatment process: activated sludge (AS), before disinfection (BD) and after disinfection (AD). Grab effluent samples were collected for bacterial isolation from each location. All samples were mixed and stored on ice, then transported to the laboratory for further processing.

Table 2.1. Process characteristics for East Lansing and Traverse City wastewater treatment plants.

	East Lansing WWTP	Traverse City WWTP
Wastewater treatment process (biological treatment)	Conventional Activated Sludge (CAS)	Membrane Biological Reactor (MBR)
Sludge Retention Time (SRT)	14 days	7.58 days
Capacity	18.8 MGD	17.0 MGD
Average Flow	13.4 MGD	8.5 MGD
Discharge Rate	14.1 MGD	4.0 MGD
Disinfection	Chlorine (Cl)	Ultraviolet (UV)

2.2. Sample Processing and Filtration

Bacteria in the effluent samples were concentrated by filtration with 0.45 μm HA filters (Millipore, Billerica, MA). The volume of AD and BD samples filtered was 1 L for each of the four samples. The filters were collected in sterile 50 mL polypropylene tubes and 50 mL Phosphate Buffer (1X PBS) was added in each tube containing a filter. The tubes were vortexed for five minutes to allow the biomass layer on the filters to mix with water. 50 mL of the AS samples were also collected in sterile centrifuge tubes. All tubes were centrifuged for 20 min at 4500 rpm to concentrate the sample down to 2 mL. Supernatant was discarded and the concentrates were stored at −80 °C until DNA extraction was performed.

2.3. Nucleic Acid Extraction

Bacterial DNA was extracted using a MagNA Pure Compact DNA extractor (Roche Applied Science, Indianapolis, IN, USA) following the protocol in the manufacturer's manual. The MagNA Pure Compact utilizes a magnetic-bead technology for the isolation process. Sample amount of 400 μL was loaded in the system and the elution volume was 100 μL. The purified DNAs were stored in a freezer at -20°C. DNA concentration was determined using the NanoDrop Spectrophotometer (NanoDrop® ND-1000, Wilmington, DE).

2.4. High-throughput Sequencing and Preprocessing

All six bacterial DNA samples were isolated and approximately 1 μg of DNA (per sample) was sent to the Research Technology Support Facility (RTSF) at Michigan State University. The NuGEN Ovation Ultralow Library System, with an input requirement of 1-100 ng of DNA, was used for all samples to accommodate for any sample containing low genetic material. After preparation, all libraries were sequenced on an Illumina platform (Illumina HiSeq2500, Roche Technologies) generating 150 bp paired-end reads.

The reads were returned as FASTQ.GZ files. All FASTQ.GZ files were processed using a Unix/Linux system offered through the MSU High Performance Computing Center (HPCC). All raw sequence reads were analyzed for quality using FastQC, a quality control tool for sequencing data (Andrews, 2010). Based on the quality control check, Illumina adapters and low quality bases were removed from all raw reads using Trimmomatic (Bolger et al., 2014). Finally, FastQC was performed once more to ensure the integrity of the sequence reads and the accuracy of the latter genetic alignment processes.

2.5. Sequence-based Metagenomic Alignment and Calculations

All six metagenomes were analyzed using Bowtie2 and Burrow-Wheeler Aligner-Maximal Exact Match (BWA-MEM), tools for aligning sequence reads to reference genomes. Bowtie2 utilizes the full-text minute (FM) index, based on the Burrows-Wheeler transform (BWT), which allows for gapped-read alignment (Langmead and Salzberg, 2012). Bowtie2 was run using default settings (end-to-end alignment, and a minimum threshold alignment score of -90) for each metagenome (Langmead and Salzberg, 2012). BWA-MEM, similar to Bowtie2, uses the FM index and allows for long gapped-read alignments. BWA-MEM was run using default settings (local alignment) (Li, 2013). A threshold of ≥ 90% gene similarity when mapped with either Bowtie2 or BWA-MEM was established in order for resistant genes to be considered in this study (Kristiansson et al., 2011; Shi et al., 2013; Zhang et al., 2015c). Resistant genes detected at ≥ 90% similarity were then clustered together for further analysis.

The breadth of coverage (gene similarity), and average depth and standard deviation (Std) were extracted for all ARGs in the reference genomes. Gene similarity, and average depth and standard deviation of all genes considered in this study can be found in the appendix section (Tables A.5 – A.10) The gene similarity for each gene was determined by normalizing the number of unique aligned bases in each position to the length of the reference gene (Molnar and Ilie, 2014). The depth, also known as the redundancy of coverage, for each gene was calculated by normalizing the total number of bases aligned in the reference gene against its total length. The redundancy of coverage for each metagenome was calculated using the Lander/Waterman equation $C = LN/G$, where L is the read length, N is the number of aligned reads and G is the haploid genome size (Sims et al., 2014).

2.6. ARG Reference Databases

Reference genomes from the Comprehensive Antibiotic Resistance Database (CARD) and the Antibiotic Resistance Gene-ANNOTation (ARG-ANNOT) database were downloaded and used for alignment. Both databases are composed of antibiotic resistant gene nucleotide sequences in FASTA format and consist of various antibiotic classes. The CARD and ARG-ANNOT references are 2,848,122 bp and 1,463,892 bp, respectively and consist of genetic data imported from NCBI GenBank and peer-reviewed publications (Gupta et al., 2015; McArthur et al., 2013). CARD and ARG-ANNOT sequences are classified based on the CARD NCBI taxonomy ontology and PubMed publications (McArthur et al., 2013), and a unified nomenclature (Gupta et al., 2015), respectively. All sequences were curated prior to being introduced into the databases. Statistical data (mapped reads, error rate, mismatches, insertions, and overall mapping quality) was retrieved directly from the sample's Binary Alignment/Mapping (BAM) output file using SAMtools.

2.7. Statistical Assessment of Sequence Alignment

Mapping quality **(**MAPQ) scores per sequence were extracted from the Sequence Alignment/Mapping (SAM) files generated from Bowtie2 and BWA-MEM alignments. The AD samples from the CAS and MBR WWTPs were considered in the statistical analysis to generate a base-level assessment of the sequence aligner's mapping performance. All sequences considered were sequence reads that were aligned during primary alignment, paired, and mapped in its proper pair (i.e. among present SAM flags, these were: 83, 99, 147, and 163). The probability of incorrect sequence alignment was calculated using the Phred scale (Ruffalo et al., 2012).

3. RESULTS

3.1. Preprocessing Quality Analysis

Quality analysis results on raw reads yielded an average guanine-cytosine (GC) content and mean quality of 56.2% and 36 for all reads, respectively. The total number of sequences per paired-end read for all samples range between 8.9-11.7 Mbp with an average of 10.4 Mbp (Table 3.1).

Table 3.1. Quality control analysis results on raw sequence reads using FastQC

Sample	Sample Abbreviation	Sequences Per Pair (bp)	Average GC Content (%)
Conventional Activated Sludge _ Activated Sludge	CAS_AS	9994682	56
Conventional Activated Sludge _ Before Disinfection	CAS_BD	10129270	55
Conventional Activated Sludge _ After Disinfection	CAS_AD	11667426	53
Membrane Bioreactor _ Activated Sludge	MBR_AS	11042222	57
Membrane Bioreactor _ Before Disinfection	MBR_BD	8879956	58
Membrane Bioreactor _ After Disinfection	MBR_AD	10670819	58

3.2. Bowtie2 and BWA-MEM alignment on Illumina high-throughput sequences

The average error rate, and GC content of mapped reads was 2.49% and 2.42%, and 53.18% and 47.8%, for CARD and ARG-ANNOT reference databases, respectively when aligned with Bowtie2. BWA-MEM generated an average error rate of 5.44% and 1.80%, and average GC content of 51.11% and 52.67% for CARD and ARG-ANNOT databases, respectively (See Tables A.1-A.4 for alignment statistics per sample). The average and maximum read length for BWA-MEM local alignment was 149 bp and 150 bp, respectively for all samples, making it analogous to Bowtie2 end-to-end read alignment of 150 bp. The number of reads mapped with BWA-MEM was 5.98 and 5.12 times greater than Bowtie2 alignment against the CARD and ARG-ANNOT reference genomes, respectively.

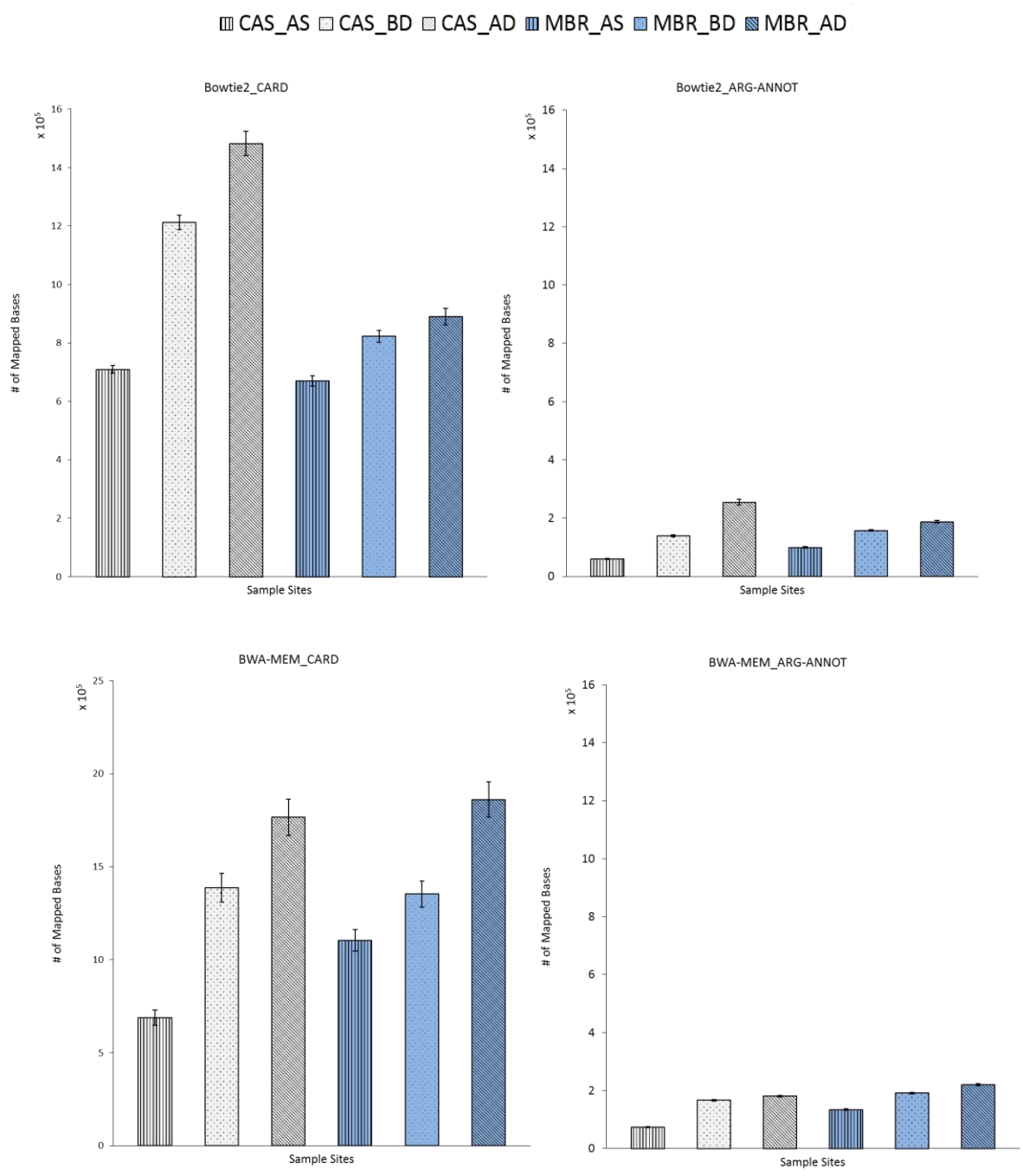

Figure 1. Number of mapped bases in CARD and ARG-ANNOT reference databases per sample. The number of mapped bases was pulled from the standard CIGAR string from SAMtools statistical computation. Error bars indicate the number of mismatches computed given the corresponding BAM file.

The number of bases mapped per sample for each reference genome reported the lowest alignment in AS samples for each treatment utility with AD samples containing the greatest

number of mapped reads (Figure 1). BWA-MEM covered 49.5% and 26.29% of ARGs for CARD and ARG-ANNOT reference genomes, respectively. This was about 14.8 and 10.2 times greater than Bowtie2. However, a significant portion of detected ARGs revealed less than 10% gene similarity, 70.61% and 85.57% for CARD and ARG-ANNOT databases, respectively. The average depth per sample for each database during BWA-MEM alignment closely corresponded to the average depth obtained during Bowtie2 alignment (Figure 2).

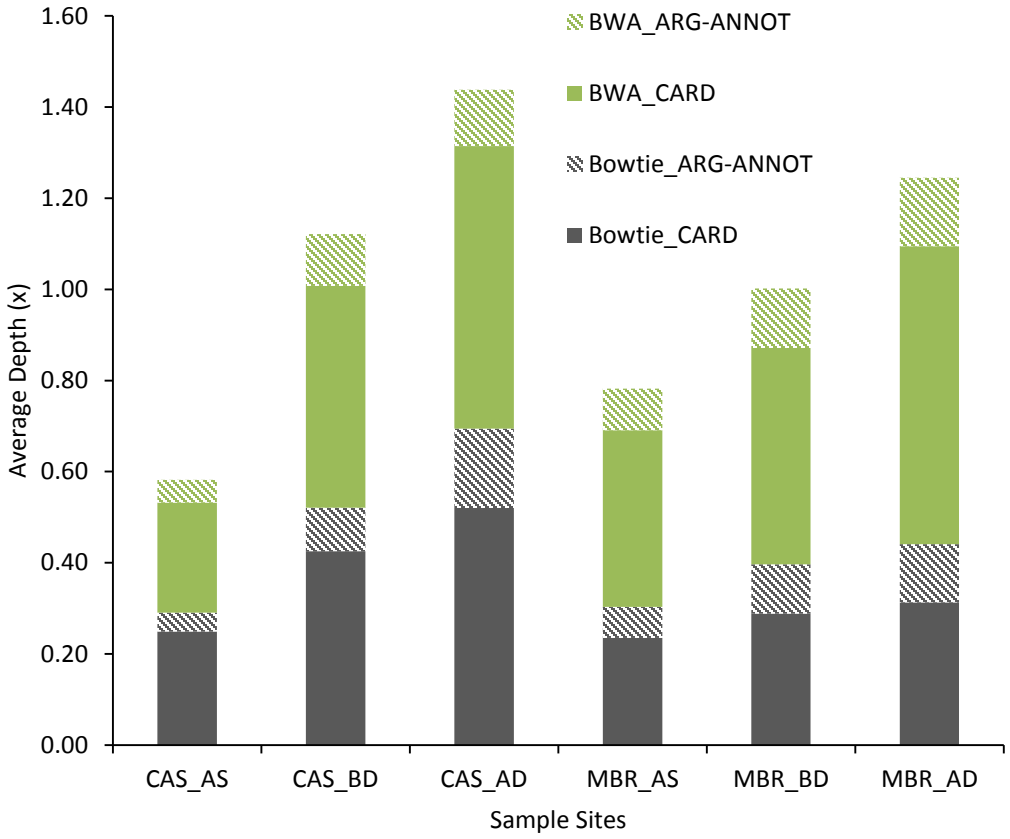

Figure 2. Comparison of average sequencing depth per sample when aligned with CARD and ARG-ANNOT nucleotide reference genomes for Bowtie2 and BWA-MEM.

3.3. ARGs Profiles in Wastewater Treatment Plants

Bowtie2 alignment of CAS samples against the CARD database yielded greater alignment per base compared to the MBR WWTP by 1.5 and 1.7-fold for BD, and AD samples, respectively.

A 1.1-fold difference was found between AS samples in CAS and MBR WWTPs. The ARG-ANNOT database found a greater number of aligned bases in AS and BD samples for the MBR plant by 1.7, and 1.1-fold, respectively. Similar to the CARD database, CAS samples yielded greater alignment in the AD sample by 1.4-fold compared to the MBR when aligned against the ARG-ANNOT database (Figure 1). BWA-MEM alignment detected a greater number of mapped bases in the MBR samples relative to the CAS samples for AS and AD sites during alignment against the CARD database. (Figure 1). Overall, results reveal a systematic trend of increasing ARGs as treatment progresses AD>BD>AS (Figure 1).

3.4. ARG Nucleotide Reference Databases

Alignment against the ARG-ANNOT reference database mapped a significantly lower number of sequence reads relative to CARD, approximately, 6.4 and 9.5 times less for Bowtie2 and BWA alignments, respectively. However, the lower number of mapped reads in the ARG-ANNOT reference database is expected considering its lesser genome size of 1463892 bp containing 1691 genes, compared to 2848122 bp with 3008 unique ARGs. ARG-ANNOT database detected only a small portion of ARGs with a gene identity of ≥ 90% in all samples given both aligners. Thus, the majority of prevalent ARGs found in all samples were identified when aligned to the CARD database. Genes resistant to Sulfonamide (*sulI*) (Genbank accession no. AF071413), detected in AS, BD, and AD samples, macrolides (*msr(E), mph(E)*) (Genbank accession no. JF769133), detected in AS, BD, and AD samples, and streptomycin (*strA*) (Genbank accession no. AB366441), detected in BD samples, with greater than 90% coverage were detected in ARG-ANNOT reference database. Tetracycline resistant gene *(tet(39)*) (Genbank accession no. AY743590), detected in BD and AD samples, was detected in both databases with identical gene

similarities. The remaining ARGs considered in latter discussions were derived from the CARD reference database.

3.5. Occurrence of ARGs in Wastewater Treatment Plants

A total of 23 different resistant genes were detected at ≥ 90% gene similarity to its reference gene and therefore, were clustered together and considered in this study. The profiles of prevalent antibiotic resistant genes in each sample resemble the profiles of the total number of mapped bases presented in Figure 1, see Figure 3.

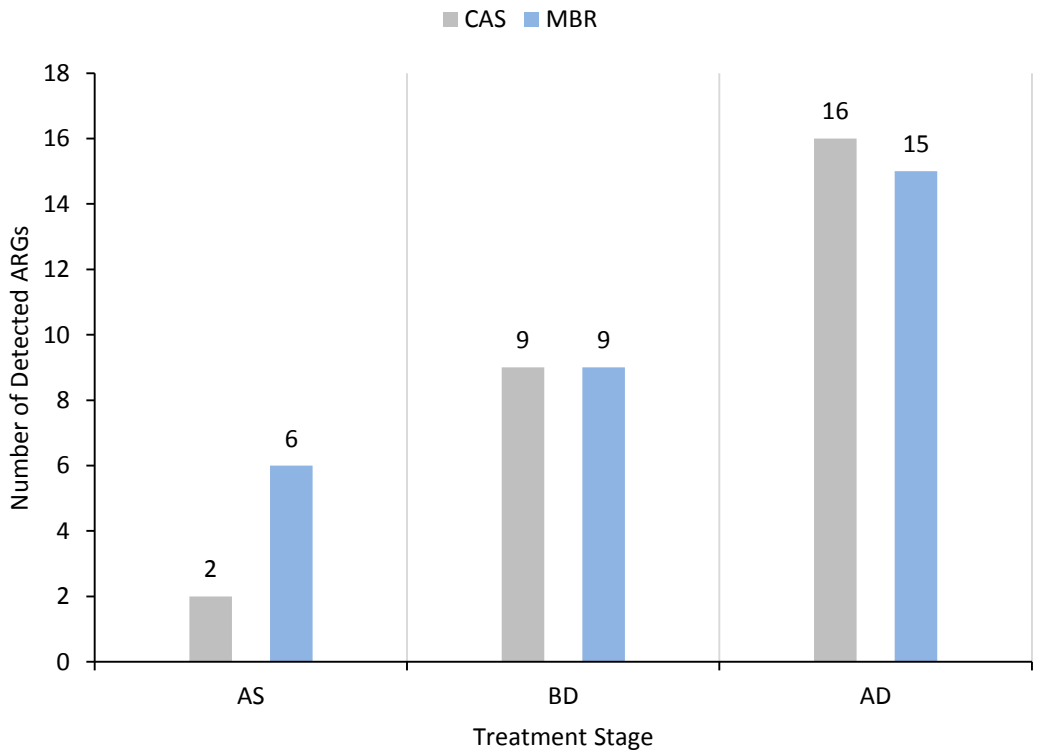

Figure 3. Profile of ARGs detected at ≥ 90% gene similarity when aligned with either Bowtie2 or BWA-MEM in each WWTP. ARGs considered consist of genes from both reference databases.

Genes with resistance to macrolides, sulfonamides, elfamycins, β-lactam, and aminoglycoside classes of antibiotics were predominant, with AD samples containing the highest occurrences of these ARGs (Figures 3 and 4).

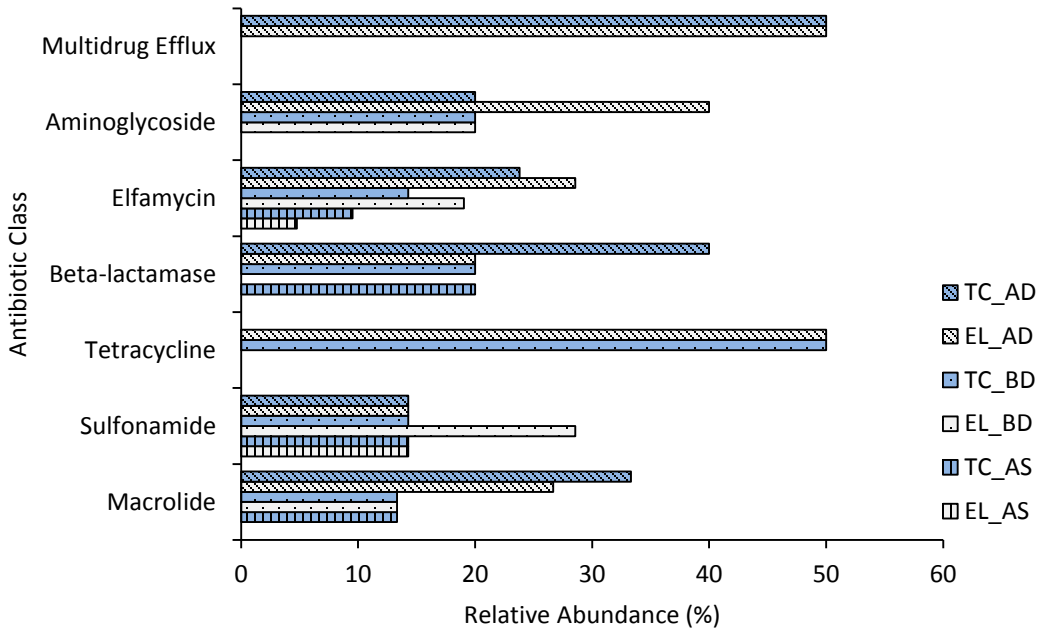

Figure 4. Relative abundance of antibiotic classes per sample belonging to resistant genes acquiring ≥ 90% gene similarity when aligned with either Bowtie2 of BWA-MEM. Genes were obtained from both CARD and ARG-ANNOT databases.

Metagenomic alignment found sulfonamides, macrolides, β-lactam, and elfamycin ARGs to be predominant in AS samples (Table 3.2). *SulI* (GenBank accession no. AF071413) was detected with Bowtie2 alignment at a 98% and 96% similarity to the reference gene for CAS and MBR WWTPs, respectively. Elfamycin resistant (*tuf)* (Genbank accession no. CP002695) genes revealed 99% and 100% gene similarity when aligned with the MBR samples according to BWA-MEM alignment. β-lactamase (bla_{RTG-5}), and erythromycin (*ermB*) resistant genes conferred a 100% gene similarity with the MBR sequence reads for AS when aligned with BWA-MEM (GenBank accession nos. JQ364968 and X64695, respectively).

Tet39, sul1, msr(E), mph(E), streptomycin (*strA, rpsL*), and *tuf* resistant genes were prevalent in BD samples (Table 3.2). Macrolides and *sulI* resistant genes possessed ≥ 93% coverage in both WWTPs (Tables A.7-A.8). *Tet(39)* was detected in the MBR WWTP with a gene identity of 94.6% and 100% from Bowtie2 and BWA alignments, respectively. *Tuf* genes obtained

a gene homology as high as 100% when aligned with BWA-MEM for both CAS and MBR BD samples (Tables A.7-A.8).

The greatest number of prevalent ARGs during Bowtie2 and BWA-MEM alignment occurred in AD samples (Figure 3). Significant portions of ARGs were not detected until AD samples in either Bowtie2 or BWA alignment. A total of 16 prevalent ARGs were detected in CAS samples, of which 56.25% of them arose after disinfection. MBR samples detected 15 ARGs, of which 53.33% of them were detected after disinfection (Figure 3 and Table 3.2). The portion of genes that persisted in AD samples from the activated sludge process was 12.50% and 33.33% for CAS and MBR treatment plants, respectively. *Bla_{RTG-5}*, *tet39*, *sul1*, aminoglycoside (*aadA6*, *msr(E)*, *mph(E)*, *tuf*, *rpsL and multidrug resistant efflux transporter (mexF)* ARGs were found to be prominent in both AD samples (Table 3.2).

Table 3.2. ARGs and corresponding ARB detected with ≥ 90% gene similarity when aligned with either Bowtie2 or BWA-MEM in AS, BD, and AD samples from CAS and MBR WWTPs. Contains genes detected in CARD and ARG-ANNOT reference databases. All annotations are generated from their respective reference database. Detected at ≥ 90% similiality (+), detected at < 90% similarity (-).

Database Accession No.	Best Hit NCBI Genbank Accession No.	Gene	ARG Class	ARB	CAS AS	MBR AS	CAS BD	MBRB D	CAS AD	MBR AD
AM087411.1.gene3	AM087411	*aadA6*	Aminoglycoside	Pseudomonas aeruginosa	-	-	-	-	+	+
JQ364968.1.gene6	JQ364968	*bla*$_{RTG-5}$	Beta-lactamase	Acinetobacter baumannii	-	+	-	+	+	+
X58272.1.gene1	X58272	*bla*$_{OXA-5}$	Beta-lactamase	Pseudomonas aeruginosa	-	-	-	-	-	+
X64695.1.gene9	X64695	*ermB*	Macrolides	Streptococcus pyogenes	-	+	-	-	-	-
NC_002516.2.882884	NC_002516	*mexF*	Multidrug Efflux	Pseudomonas aeruginosa	-	-	-	-	+	+
(MLS)MphE:JF769133:8777-9661:885	JF769133	*mph(E)*	Macrolides	Pasteurella multocida	-	+	+	+	+	+
(MLS)MsrE:JF769133:7246-8721:1476	JF769133	*msr(E)_1*	Macrolides	Pasteurella multocida	-	-	+	+	+	+
EF102240.1.gene3	EF102240	*msr(E)_2*	Macrolides	Acinetobacter baumannii	-	-	-	-	+	-
NC_007682.3.4246769	NC_007682	*msr(E)_3*	Macrolides	Escherichia coli	-	-	-	-	-	+
NC_010481.6155782	NC_010481	*msr(E)_4*	Macrolides	Acinetobacter baumannii	-	-	-	-	+	+
CP003022.1.gene337	CP003022	*msr(E)_5*	Macrolides	Pasteurella multocida	-	-	-	-	-	+
AL123456.3.gene715	AL123456	*rpsL_1*	Aminoglycoside	Mycobacterium tuberculosis	-	-	-	+	-	-
CP003248.2.gene711	CP003248	*rpsL_2*	Aminoglycoside	Mycobacterium tuberculosis	-	-	-	+	-	+
(AGly)StrA:AB366441:22458-23261:804	AB366441	*strA*	Aminoglycoside	Salmonella enterica	-	-	+	-	-	-
(Sul)SulI:AF071413:6700-7539:840	AF071413	*sulI_1*	Sulfonamide	Escherichia coli	+	+	+	+	+	+
AJ223604.1.gene9	AJ223604	*sulI_2*	Sulfonamide	Pseudomonas aeruginosa	-	-	+	-	-	-
AY743590.gene	AY743590	*tet(39)*	Tetracycline	Acinetobacter LUH5605	-	-	-	+	+	-
CP002695.1.gene3614	CP002695	*tuf_1*	Elfamycins	Bordetella pertussis	+	+	+	+	+	+
NC_011595.7072242	NC_011595	*tuf_2*	Elfamycins	Acinetobacter baumannii	-	-	+	+	+	+

Table 3.2. (cont'd)

CP002695.1.gene10	CP002695	*tuf_3*	Elfamycins	Bordetella pertussis	-	+	+	+	+	+
CP000647.1.gene3761	CP000647	*tuf_4*	Elfamycins	Klebsiella pneumoniae	-	-	-	-	+	-
NC_002516.2.881718	NC_002516	*tufA*	Elfamycins	Pseudomonas aeruginosa	-	-	-	-	+	+
NC_002516.2.881697	NC_002516	*tufB*	Elfamycins	Pseudomonas aeruginosa	-	-	+	-	+	+

E.coli, *Acinetobacter baumannii*, *Pasteurella multocida*, and *Pseudomonas aeruginosa* composed the majority of antibiotic resistant bacteria (ARB) in AD samples (Table 3.2). Resistant gene *msr(E)_2* belonging to *Acinetobacter baumannii* possessed a 97% gene identify in the AD sample for CAS treatment plant during Bowtie2 alignment, but was not detected in the MBR sample. Sulfonamide resistant gene, *sul1*, belonging to *E.coli* conferred a 98% (97%) and 98% (98%) gene identity for CAS and MBR after disinfection samples, respectively (Table A.9 – A.10).

Beta-lactamase (RTG-5) belonging to *Acinetobacter baumannii* conferred a 99% gene identity in AD samples for the CAS treatment plant with an average depth of 1.27 and 1.36x in Bowtie2 and BWA-MEM alignment processes, respectively. While $bla_{RTG\text{-}5}$ resistant gene went undetected in the MBR AD sample during Bowtie2 alignment, a 94% similarity was found during alignment with BWA-MEM. *Pasteurella multocida* resistant gene, *mph(E),* possessed a 98% (93%) and 91% (98%) identity to samples from the CAS and MBR AD samples, respectively, during alignment with Bowtie2 (BWA-MEM). After alignment with BWA-MEM, AD samples from CAS and MBR conferred a 100% gene identity with an average depth of 3.51 and 2.56, respectively to *aadA6* resistant gene belonging to *Pseudomonas aeruginosa* (Table A.9 – A.10).

3.6. Statistical Assessment using MAPQ Scores generated from Bowtie2 and BWA-MEM Alignments

MAPQ scores were extracted from Bowtie2 and BWA alignment from their respective SAM file. BWA-MEM produced MAPQ scores ranging from 0 to 60 ($p<0.001$). Average MAPQ scores generated during alignment with the CARD database were 18.65 ($p=0.01$) and 15.76 ($p=0.03$) for CAS and MBR samples, respectively. Alignment with the ARG-ANNOT databases revealed an average MAPQ score of 55.89 ($p<0.001$) and 59 ($p<0.001$) for CAS and MBR samples, respectively.

The lowest score reported for Bowtie2 alignment was 0 ($p=1$) with the highest being 255 (MAPQ score not available). The MAPQ score of selected sequences in Bowtie2 was reported as 255 for the majority of unique reads (reads with only one sequence alignment reported). Multi-reads (reads with multiple alignments) reported in Bowtie2 generally contained a score of either 0 or 1 ($p>0.5$) depending on the alignment score of the best alignment, and the difference between the best and second best alignment scores. Due to the absence of usable MAPQ scores for unique reads for Bowtie2, its base-line sequence mapping quality could not be fully assessed in this study.

4. DISCUSSION

4.1. Occurrence of ARGs in WWTPs

Here, we use a sequence-based metagenomic approach to investigate the prevalence of a broad spectrum of ARGs in a conventional activated sludge WWTP with chlorine disinfection, and membrane bioreactor with UV disinfection. In this study, genes resistant to tetracycline, sulfonamide, β-lactam, aminoglycoside, elfamycin, and macrolide classes of antibiotics along with a multidrug efflux transporter (*mexF)* were detected in both CAS and MBR WWTPs.

Tetracycline and sulfonamide resistant genes are among the most common ARGs found in wastewater treatment systems, and other aquatic environments due to their widespread use (Gao et al., 2012a; Mao et al., 2015; Zhang et al., 2009a; Zhang et al., 2015a). *Tet(39),* and *sulI* were detected in BD and AD samples. Tetracycline and sulfonamide resistant genes are known to persist during wastewater treatment (Mao et al., 2015; Szczepanowski et al., 2009; Vaz-Moreira et al., 2015). The ability of these genes to withstand various treatment processes greatly increases the risks of these genes contaminating natural water bodies. Mao et al., (2015) evaluated the profile of 30 different ARGs in WWTPs in China and revealed that *tet* and *sul1* genes were enriched in the final effluent of both selected WWTPs. Another study, conducted between hospital effluent, and the influent and effluent of an urban WWTP, revealed higher concentrations of sulfonamide resistant genes in the treated wastewater and hospital effluent compared to the raw sewage entering the WWTP (Vaz-Moreira et al., 2015). In contrast, Munir et al., 2011 observed a reduction in *tetO*, *tetW*, and *sul1* resistant genes in the effluent of five different WWTPs in Michigan containing various treatment types. Another study, conducted on tetracycline resistant genes, *tetA* and *tetC,* in different water environments, found that the concentration of both tetracycline resistant genes were reduced in the final effluent of a sewage treatment plant in China (Zhang et al., 2009a).

Theses discrepancies indicate that antibiotic resistance selection is dependent on factors beyond that of solely wastewater treatment.

Class A, RTG-5 (CARB-14), beta-lactamase, conferring resistance to carbenicillin (Bonnin et al., 2012; Couture et al., 1992), was detected in all MBR samples and in the AD sample of the CAS WWTP. Class D, OXA-5, beta-lactamase, conferring resistance to oxacillin (Couture et al., 1992), was detected in the MBR AD sample. A primary mechanism behind β-lactamase resistance, generally for gram-negative bacteria, is the production of β-lactamase enzymes (Couture et al., 1992; Xia et al., 2016). Beta-lactams are a widely used antibiotic with wide occurrences of resistance within the microbial community due to its extensive use and low toxicity (Zhang et al., 2009b). A comprehensive study on bacterial plasmids isolated from the activated sludge and final effluent of a WWTP in Germany also detected β-lactamase hydrolyzing enzyme, OXA-5 (Genbank accession no. JQ364968) present in AS and final effluent samples as well as resistance to several other class A-D β-lactamases (Szczepanowski et al., 2009).

Multidrug efflux transporter, *mexF*, was detected in AD samples of both CAS and MBR treatment plants. The multidrug resistance pump, *mexF*, is a part of the resistance-nodulation-cell division (RND) family, encoding resistance in *Pseudomonas aeruginosa* to antibiotics such as, chloramphenicol, trimethoprim, and fluoroquinolones. The efflux pump is said to consist of three proteins, which collaborate to expel antibiotics from the cytoplasm or periplasm of the microorganism (Aires et al., 2002). A number of studies have reported RND resistance pumps in aquatic environments (Gomez-Alvarez et al., 2012; Szczepanowski et al., 2009).

Aminoglycosides, *aadA6* and *rpsL*, conferring resistance to streptomycin and spectinomycin (Finken et al., 1993; Papadovasilaki et al., 2015) were also present in this study. Aminoglycosides

consist of three groups, phosphotransferases (APHs), acetyltransferases (AACs), and the adenyltransferases (ANTs) (Wright, 1999). They are used against a wide range of aerobic bacteria, including *Mycobacterium tuberculosis* (Fonseca et al., 2015). Streptomycin was the first antibiotic used to treat TB (Honore et al., 1994). *M.tuberculosis* bacteria resistant to this drug are often classified as extensively drug-resistant strains (Fonseca et al., 2015). Mutations in the ribosomal protein S12 encoding gene, *rpsL*, is said to cause resistance to streptomycin in *M.tuberculosis* (Finken et al., 1993). Aminoglycoside resistant genes have become increasingly common in water environments (Shi et al., 2013; Szczepanowski et al., 2009) and have been reported in the effluent of treated wastewater in several instances (Szczepanowski et al., 2009; Vaz-Moreira et al., 2015).

Macrolide and elfamycin resistant genes were the most abundant genes detected in this study with resistant genes occurring in all six samples. After disinfection samples produced the greatest abundance of these genes compared to AS and BD samples. Genes resistant to macrolide classes of antibiotics composed 30.43% (7 out of 23) of all prevalent ARGs detected in this study. Macrolide resistant genes *msr(E)* and *mph(E)*, detected in this study, encode a macrolide efflux pump, and macrolide-inactivating phosphotransferase, respectively (Rose et al., 2012). Rose et al., (2012) revealed that *msr(E)* and *mph(E)* were consistently expressed in tandem when detected in *Pasteurella multocida*, which is revealed in one instance in our study. Erythromycin resistant gene *ermB*, belonging to macrolide resistant classes of antibiotics, was detected in the AS sample from the MBR WWTP. A recent study, conducted on a bench-scale activated sludge treatment process, observed a survival rate of 100% among bacteria that were highly resistance to erythromycin in the effluent of the activated sludge process. It also revealed that the proportions of *ermB* remained constant in the activated sludge process (Guo et al., 2015). In addition to this, erythromycin

resistant genes are known to persist in the final effluent of WWTPs (Berglund et al., 2015; Sidrach-Cardona et al., 2014).

Similar to macrolide ARGs, elfamycin resistant genes were the second most prevalent genes detected in this study composing 26.09% (6 out of 23) of ARGs detected. Elfamycins are a class of naturally occurring antibiotics, consisting of antibiotics such as Kirromycin, Aurodox, and Efrotomycin (Hall et al., 1989; Sottani et al., 1993; Miele et al., 1994). Elfamycins inhibit bacterial growth by binding to the elongation factor Tu polypeptide, a component responsible for bacterial protein synthesis (Sottani et al., 1993). To our knowledge, no recent studies on the prevalence of elfamycin resistant genes in sewage treatment plants have been documented. Sequence-based metagenomics has revealed the occurrence of these genes in all treatment stages with reports of 100% gene similarity to the reference gene.

4.2. Occurrence of ARGs in Various Wastewater Treatment Stages

Activated sludge samples from both WWTPs attained the lowest occurrence of ARGs compared to before disinfection and after disinfection samples with no notable difference between the two biological treatment processes. It has been reported that biological treatment types, loading rates, sludge retention times and the production of biosolids, to name a few, impact the presence of resistant genes (Kim et al., 2010; Munir et al., 2011; Novo and Manaia, 2010) in biological treatment processes. Kim et al. (2010) reported the occurrence of tetracycline resistant genes in three WWTPs located in New York and Connecticut, with varying biological treatment functions. The study revealed that the AS processes did not contribute significantly to the enrichment of tetracycline resistant genes, and found that the fraction of these genes were lower in AS processes compared to the before disinfection processes. This was consistent with our study, which revealed a systematic trend of increasing occurrence of ARGs as treatment progressed.

The low profile of ARGs in the AS process could be the result of several factors such as, nutrient abundance, and adsorption to sludge biosolids. Munir et al. (2011) reported an increase in the concentration of *tet* and *sul* ARGs studied in the biosolids of five different WWTPs by several orders of magnitude relative to before and after disinfection samples. These ARGs can potentially be exposed to the natural environment via land applications causing further exposure and spread of antibiotic resistance (Schmieder and Edwards 2012). A study conducted by Novo and Manaia (2010) documented the prevalence of ARGs in three different WWTPs with varying biological treatment processes. The study revealed that ARGs resistant to tetracycline persisted in the treated effluent of the conventional activated sludge WWTP compared to the trickling filter, and submerged aeration filter facilities. The mechanisms behind the prevalence of ARGs after biological treatment systems are still unclear. Further studies are needed to assess various operational, functional, and environmental factors that contribute to the composition of resistant genes in wastewater microbial communities.

Furthermore, a systematic increase was observed in the profile of ARGs from AS to BD samples for both WWTPs. There was an increase in genes conferring resistance to elfamycins, and the occurrence of *tet(39), strA,* and *rpsL* resistant genes in BD samples for each WWTP. Similar to our results, other studies have reported an increase in antibiotic resistant genes like tetracycline and macrolides in before disinfection samples (Guo et al., 2015; Kim et al. 2010) suggesting that resistant genes arose either in the effluent of the biological treatment processes, or during secondary clarification.

After disinfection samples revealed the largest occurrence of ARGs in both CAS and MBR WWTPs. A significant number of the 23 total ARGs, which were detected at ≥ 90% gene

similarity, were not detected until after disinfection in both CAS and MBR treatment plants, 56.25% and 53.33%, respectively. Consistent with recent studies, a number of disinfection processes fail to be effective against reducing ARGs, and in some cases, an increase in the occurrence of these genes was observed in the treated effluent (Huang et al., 2011; Murray et al., 1984; Shi et al., 2013; Yuan et al., 2015). A previous study revealed a significant increase in ARGs in water sampled from a chlorine disinfection tank as opposed to the raw source water of a water treatment plant in Nanjing, China. The results also revealed a significant increase in mobile genetic elements in the chlorinated water compared to the raw water sources (Shi et al., 2013). Other studies reported that chlorine disinfection had little effect on reducing the presence of ARGs in the effluent of WWTPs (Yuan et al., 2015; Munir et al., 2011). On the other hand, several studies have reported a decrease in the presence of ARGs after chlorination (Huang et al., 2011; Mao et al., 2015; Murray et al., 1984).

Similar to chlorine disinfection, recent studies detected an increase in antibiotic resistance after UV disinfection (Kim et al., 2010). In our study, there was an increase in detected ARGs after UV radiation treatment, but there was no significant difference between the occurrences of ARGs following chlorination compared to UV.

4.3. Composition of ARBs in Chlorine and UV Treated Wastewater

Acinetobacter baumannii, *Pseudomonas aeruginosa*, and *Pasteurella multocida* composed the majority of antibiotic resistant bacteria found in AD samples in both WWTPs. According to the CARD and ARG-ANNOT resistance gene ontologies (ARO), a number of prevalent ARGs found in this study: including macrolides, elfamyins, multidrug, aminoglycoside and β-lactamase resistant genes, belong generally to the abovementioned gram-negative bacteria. *A. baumannii* is an opportunist pathogen responsible for a number of nosocomial infections and reported outbreaks

within recent years (Zarrilli et al., 2008). Since the rise of its antibiotic resistant characteristics in the 1960s, *A. baumannii* has since been characterized as a multidrug resist bacteria known to develop resistance within the first year of introducing a new antibiotic for treatment (Gonzalez-Villoria and Valverde-Garduno, 2016).

Similar to *A. baumannii*, *Pseudomonas aeruginosa* is an opportunistic hospital-acquired pathogen commonly found in moist, high nutrient environments such as WWTPs (Slekovec et al., 2012). *P. aeruginosa* accounts for 15% of hospital acquired infections to date, mainly attacking intensive care units. These pathogenic bacteria also play a role in community-acquired infections due to exposure to contaminated water sources (Slekovec et al., 2012). Slekovec et al., 2012 found the presence of *P. aeruginosa* to be most abundant in water extracted from the effluent of a hospital waste stream and the treated sludge from a WWTP, than the other environments sampled. Furthermore, the study revealed the appearance of multidrug resistant strains of *P.aeruginosa* in the hospital wastewater, treated wastewater, and receiving river water downstream the WWTP. The occurrence of these types of resistant opportunistic pathogens in the environment is of increasing importance concerning human health (Bouki et al., 2013; Varela and Manaia, 2013).

Pasteurella multocida was also reported in this study and persisted in each of the treatment stages in the MBR WWTP. *P. multocida* is a pathogenic bacterium occurring in the natural oral flora of several species of animals (Weber et al., 1984). *P. multocida* are commonly classified as zoonotic pathogens known to cause various infectious diseases in both animals and humans (Rose et al., 2012; Weber et al., 1984). Human infections usually occur as a result of an animal bite causing a number of diseases including meningitis and pneumonia (Weber et al., 1984). Reports

have classified several antibodies that are no longer effective against *P. multocida* due to its reduced susceptibility to the drug (Rose et al., 2012; Sellyei et al., 2009; Weber et al., 1984).

Other bacteria such as *E.coli, Bordetella pertussis, Klebsiella pneumonia*, and streptomycin resistant *Mycobacterium tuberculosis* were also reported in AD samples.

Several studies reveal significant concentrations of antimicrobial resistant pathogenic bacteria in treated wastewater despite reductions of these bacteria throughout the treatment process (Everage et al., 2014; Munir et al., 2011; Slekovec et al., 2012). These studies indicate that resistant bacteria are disseminating into natural water bodies despite treatment (Gao et al., 2012a).

4.4. Statistical Analysis of Sequence Alignment using MAPQ Scores

Genetic alignment of AS, BD, and AD samples from each treatment utility was performed using Bowtie2 and BWA-MEM, which have proven to be accurate sequence aligners (Langmead and Salzberg, 2012; Li, 2013). Sequence-based metagenomics is most accurate when identifying known genes (Schmieder and Edwards, 2012), therefore ARGs found in all samples using metagenomic analysis only highly suggest the presence of these ARGs in this study. ARGs with a gene similarity less than 100% should be interpreted with caution as they could belong to some variation of the known reference gene, or novel resistance determinant (Schmieder and Edwards, 2012). Despite vast enhancements in the efficiency and accuracy of genetic aligners, they still have their limitations. Genetic alignment is exposed to a number of performance errors. For example, insertion or deletions (indels) generated during sequencing runs, as well as duplicate regions in the reference genome can lead to incorrect mapping during the alignment process (Li and Homer, 2010; Olson et al., 2015). Most sequence aligners generate sequence mapping quality scores for users to gage the confidence level of aligned sequences (Li and Homer, 2010; Ruffalo et al., 2011).

This study provided a baseline assessment of the unaltered sequence alignment errors rates from each aligner's MAPQ scores when operating under default conditions. It was found that BWA-MEM produced more useful MAPQ scores compared to Bowtie2 (Ruffalo et al., 2011) and the number of sequences passing threshold requirements was greatest in BWA-MEM when aligned with either reference database for AD samples.

Bowtie2 MAPQ scores for high quality, unique-read alignments were not accounted for when aligned against either reference database, while multi-read alignments generally received a score of either 0 or 1 depending on Bowtie2's scoring criteria found in the user's manual (Langmead and Salzberg, 2012). Thus, MAPQ scores for Bowtie2 could not be assessed without further analysis. The integrity of MAPQ scores is not only important in evaluating the accuracy of genetic alignment, but also for downstream applications such as variant analysis (Li and Homer, 2010; Olson et al., 2015). Therefore, it is a best practice to realign MAPQ scores for accuracy before proceeding to downstream applications (Olson et al., 2015).

5. CONCLUSION

The rise of antibiotic resistance genes in the environment stimulated the need to study the development of these genes in favorable environments such as WWTPs. Thus, our study aimed to explore the profile and occurrence of these genes in WWTPs using a sequenced-based metagenomic approach. This study allowed for the detection of a broad spectrum of ARGs in the activated sludge, before disinfection and after disinfection treatment stages of two WWTPs in Michigan. The results reveal a systematic increase in the number of ARGs as treatment progressed suggesting that wastewater treatment promotes the development of antimicrobial resistant genes. On the basis of conflicting studies, the selection for ARGs in WTTPs is still not fully understood. However, it is evident that a number of factors impact the selection or removal of ARGs in sewage treatment systems. Our study further contributes to the pool of information on the dissemination of ARGs in WWTPs given various operational conditions and treatment stages. Further studies, assessing a wide range of functional, operational, and environmental factors, are needed to fully understand the behavior of these genes in treatment plants and in the environment.

APPENDIX

Table A.1. Bowtie2 alignment statistics per sample

Sample	Raw Total Sequences		Total Length		Mapped Reads		% of Reads Mapped		Reads Mapped and Paired (both mates paired)		Reads Paired Properly		# of Mapped Bases (CIGAR)		Overall Mean Coverage (x)	
	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT
CAS_AS	19993744	19989583	2999061600	2998437450	4730	401	0.02	0.00	4099.00	374.00	3754.00	352.00	709500	60150	0.25	0.04
CAS_BD	20265847	20259002	3039877050	3038850300	8079	931	0.04	0.00	7090.00	778.00	6250.00	698.00	1211850	139650	0.43	0.10
CAS_AD	23343571	23335909	3501535650	3500386350	9879	1700	0.04	0.01	8605.00	1535.00	7784.00	1382.00	1481850	255000	0.52	0.17
MBR_AS	22088306	22084753	3313245900	3312712950	4466	664	0.02	0.00	3855.00	586.00	3780.00	574.00	669900	99600	0.24	0.07
MBR_BD	17764690	17760541	2664703500	2664081150	5482	1052	0.03	0.01	4786.00	970.00	4158.00	908.00	822300	157800	0.29	0.11
MBR_AD	21346764	21342423	3202014600	3201363450	5936	1253	0.03	0.01	4972.00	1134.00	4736.00	1074.00	890400	187950	0.31	0.13

Table A.2. Bowtie2 alignment statistics per sample

Sample	# of mismatches		General Error Rate		Average insert size		Avg. GC Content (%)		Upper and Lower GC Content Limits (%)			
	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ANNOT	CARD		ANNOT	
ELAS	13229	1297	1.86E-02	2.16E-02	240.8	259.9	53.82	45.35	65.56	30.58	64.43	30.06
ELBD	24418	3139	2.01E-02	2.25E-02	237.9	237.9	53.57	47.82	65.57	32.32	64.43	32.49
ELAD	41546	9307	2.80E-02	3.65E-02	244.3	242.4	53.67	49.40	67.88	33.06	64.43	32.49
TCAS	17068	2508	2.55E-02	2.52E-02	240.2	244.3	52.75	48.33	68.91	32.32	67.13	32.49
TCBD	21003	2947	2.55E-02	1.87E-02	239.6	240.5	52.43	47.40	68.19	33.20	64.43	32.49
TCAD	28224	3963	3.17E-02	2.11E-02	225.7	222.3	52.86	48.47	68.91	33.20	67.73	32.49

Table A.3. BWA-MEM alignment statistics per sample

Sample	Raw Total Sequences		Total Length		Mapped Reads		% of Reads Mapped		Reads Mapped and Paired (both mates paired)		Reads Paired Properly		# of Mapped Bases (CIGAR)		Overall Mean Coverage (x)	
	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT
ELAS	19989415	23406240	2998407467	3510934206	16725	1453	0.08	0.01	14212	1121	11823	646	688956	73346	0.24	0.05
ELBD	20258671	20258558	3038788355	3038781755	36990	3761	0.18	0.02	33995	3228	31472	2542	1386247	166824	0.49	0.11
ELAD	23334974	23334863	3500234338	3500228265	46742	3333	0.20	0.01	43390	2858	40972	2002	1765860	180636	0.62	0.12
TCAS	22084499	22084450	3312669599	3312666836	30584	3217	0.14	0.01	27805	2819	25228	2292	1103265	134197	0.39	0.09
TCBD	17759978	17759929	2663990251	2663987492	38619	5510	0.22	0.03	36028	5049	34148	4511	1353086	191266	0.48	0.13
TCAD	21341713	21341645	3201249768	3201245963	54388	6316	0.25	0.03	51408	5678	49883	5068	1861326	219762	0.65	0.15

Table A.4. BWA-MEM alignment statistics per sample

Sample	# of mismatches		General Error Rate		Average insert size		Avg. GC Content (%)		Upper and Lower GC Content Limits (%)				Average Mapping Quality	
	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD	ARG-ANNOT	CARD		ARG-ANNOT		CARD	ARG-ANNOT
ELAS	41395	1534	6.01E-02	2.45E-02	119.9	186	51.83	55.77	77.01	22.74	75.66	22.24	32.9	33
ELBD	76779	3095	5.54E-02	1.86E-02	142.2	196.3	50.49	51.64	77.01	22.74	75.66	22.24	33.2	33.2
ELAD	98055	3146	5.55E-02	1.74E-02	148	210.2	50.09	51.16	77.01	22.74	75.66	22.24	33.3	33.3
TCAS	58218	2618	5.28E-02	1.95E-02	129.7	192.1	51.86	53.12	77.01	22.74	75.66	22.24	32.8	32.8
TCBD	69102	2637	5.11E-02	1.38E-02	146.7	172.5	51.64	52.23	77.01	22.74	75.66	22.24	32.8	32.8
TCAD	95686	3180	5.14E-02	1.45E-02	155.8	175.7	50.74	52.11	77.01	22.74	75.66	22.24	32.9	32.9

Table A.5. Average depth and gene similarity with Bowtie2 alignment for ARGs with ≥ 90% coverage in activated sludge (AS) samples for East Lansing (CAS) and Traverse City (MBR) WWTPs. ARGs from both CARD and ARG-ANNOT reference database are integrated into one data set.

Database Accession No.	Gene	Length	GC%	Avg Depth (x)		Depth Std		Gene Similarity (%)	
				ELAS	TCAS	ELAS	TCAS	ELAS	TCAS
CP002695.1.gene10	*tuf*	1191	62.64	1.65	1.77	1.43	1.80	82.28	66.16
CP002695.1.gene3614	*tuf*	1191	62.64	0.52	2.09	0.50	2.31	52.98	88.58
(Sul)SulI:AF071413:6700-7539:840	*sulI*	855	60.58	7.91	6.57	3.00	4.14	98.25	96.14
(MLS)MphE:JF769133:8777-9661:885	*mphE*	900	35.78	0.40	3.65	0.49	1.47	40.78	96.33
JQ364968.1.gene6	bla_{RTG-5}	378	43.12	-	1.15	-	0.85	ND	70.37
X64695.1.gene9	*ermB*	111	31.53	-	-	-	-	ND	ND

Table A.6. Average depth and gene similarity with BWA-MEM alignment for ARGs with ≥ 90% coverage in activated sludge (AS) samples for East Lansing (CAS) and Traverse City (MBR) WWTPs. ARGs from both CARD and ARG-ANNOT reference database are integrated into one data set.

Database Accession No.	Gene	Length	GC%	Avg Depth (x)		Depth Std		Gene Similarity (%)	
				ELAS	TCAS	ELAS	TCAS	ELAS	TCAS
CP002695.1.gene10	*tuf*	1191	62.64	4.16	14.90	3.66	8.85	81.95	99.50
CP002695.1.gene3614	*tuf*	1191	62.64	5.58	15.91	3.65	7.81	96.31	100
(Sul)SulI:AF071413:6700-7539:840	*sulI*	855	60.58	0.04	0.87	0.21	0.76	4.44	74.62
(MLS)MphE:JF769133:8777-9661:885	*mphE*	900	35.78	-	1.28	-	0.77	ND	80.56
JQ364968.1.gene6	bla_{RTG-5}	378	43.12	-	2.80	-	1.66	ND	100
X64695.1.gene9	*ermB*	111	31.53	-	1.01	-	0.09	ND	100

Table A.7. Average depth and gene similarity with Bowtie2 alignment for ARGs with ≥ 90% coverage in before disinfection (BD) samples for East Lansing (CAS) and Traverse City (MBR) WWTPs. ARGs from both CARD and ARG-ANNOT reference database are integrated into one data set.

Database Accession No.	Gene	Length	GC%	Avg Depth (x)		Depth Std		Gene Similarity (%)	
				ELBD	TCBD	ELBD	TCBD	ELBD	TCBD
AY743590.gene	*tet39*	1188	40.57	1.32	2.36	1.34	0.89	64.81	94.61
(AGly)StrA:AB366441:22458-23261:804	*strA*	818	55.26	2.10	0.37	1.23	0.48	97.68	36.67
(Sul)SulI:AF071413:6700-7539:840	*sulI*	855	60.58	10.94	5.89	5.21	2.63	97.43	93.68
(MLS)MphE:JF769133:8777-9661:885	*mph(E)*	900	35.78	2.87	6.33	1.73	3.43	95.78	95.89
(MLS)MsrE:JF769133:7246-8721:1476	*msr(E)*	1501	39.31	5.05	6.31	2.67	2.91	94.47	94.74
CP002695.1.gene10	*tuf*	1191	62.64	1.27	1.74	1.29	2.32	75.06	73.97
CP002695.1.gene3614	*tuf*	1191	62.64	0.95	2.31	1.14	2.31	45.59	64.74
NC_002516.2.881697	*tuf*	1194	59.63	1.48	0.75	1.52	0.95	58.71	45.06
AJ223604.1.gene9	*sulI*	348	50	-	-	-	-	ND	ND
NC_011595.7072242	*tuf*	705	45.11	0.9872	1.18	1.10	1.54	49.08	44.54
JQ364968.1.gene6	bla_{RTG-5}	378	43.12	-	1.18	-	0.88	ND	68.25
AL123456.3.gene715	*rpsL*	375	62.67	-	-	-	-	ND	ND

Table A.8. Average depth and gene similarity with BWA-MEM alignment for ARGs with ≥ 90% coverage in before disinfection (BD) samples for East Lansing (CAS) and Traverse City (MBR) WWTPs. ARGs from both CARD and ARG-ANNOT reference database are integrated into one data set.

Database Accession No.	Gene	Length	GC%	Avg Depth (x)		Depth Std		Gene Similarity (%)	
				ELBD	TCBD	ELBD	TCBD	ELBD	TCBD
AY743590.gene	*tet39*	1188	40.57	1.16	2.26	1.21	0.92	62.79	100
(AGly)StrA:AB366441:22458-23261:804	*strA*	818	55.26	0.53	0.54	0.83	0.55	31.30	51.34
(Sul)SulI:AF071413:6700-7539:840	*sulI*	855	60.58	3.21	2.93	1.93	1.87	93.92	89.82
(MLS)MphE:JF769133:8777-9661:885	*mph(E)*	900	35.78	1.18	3.86	0.66	2.15	86.33	98.33
(MLS)MsrE:JF769133:7246-8721:1476	*msr(E)*	1501	39.31	0.78	1.97	0.99	1.40	46.50	81.81
CP002695.1.gene10	*tuf*	1191	62.64	23.59	28.30	13.96	16.49	100	100
CP002695.1.gene3614	*tuf*	1191	62.64	25.95	32.10	14.64	17.13	97.90	100
NC_002516.2.881697	*tuf*	1194	59.63	18.05	20.28	15.16	17.41	93.80	89.28
AJ223604.1.gene9	*sulI*	348	50	1.14	0.60	0.55	0.56	90.80	56.90
NC_011595.7072242	*tuf*	705	45.11	5.39	3.69	4.92	2.35	90.21	90.35
JQ364968.1.gene6	bla_{RTG-5}	378	43.12	0.47	1.95	0.50	0.45	47.62	99.74
AL123456.3.gene715	*rpsL*	375	62.67	2.37	4.61	2.85	3.43	57.87	99.20

Table A.9. Average depth and gene similarity with Bowtie2 alignment for ARGs with ≥ 90% coverage in after disinfection (AD) samples for East Lansing (CAS) and Traverse City (MBR) WWTPs. ARGs from both CARD and ARG-ANNOT reference database are integrated into one data set.

Database Accession No.	Gene	Length	GC%	Avg Depth (x)		Depth Std		Gene Similarity (%)	
				ELAD	TCAD	ELAD	TCAD	ELAD	TCAD
(Sul)SulI:AF071413:6700-7539:840	*sul1*	855	60.58	12.05	5.40	6.23	2.06	97.89	98.25
AM087411.1.gene3	*aadA6*	846	53.66	2.59	0.68	1.60	0.68	81.44	56.50
AY743590.gene	*tet39*	1188	40.57	1.23	1.85	0.59	1.17	93.60	83.08
CP002695.1.gene3614	*tuf*	1191	62.64	0.97	2.02	0.95	1.31	64.90	95.05
EF102240.1.gene3	*msrE*	1476	39.97	2.01	1.69	0.79	1.58	97.83	75.14
(MLS)MphE:JF769133:8777-9661:885	*mphE*	900	35.78	6.36	5.00	2.41	2.79	98.33	90.78
(MLS)MsrE:JF769133:7246-8721:1476	*msrE*	1501	39.31	9.53	9.06	4.26	3.57	98.33	97.14
JQ364968.1.gene6	bla_{RTG-5}	378	43.12	1.27	0.53	0.46	0.50	99.21	53.44
NC_002516.2.881697	*tufB*	1194	59.63	4.54	1.49	4.16	2.29	87.02	43.30
NC_007682.3.4246769	*msrE*	1476	39.97	2.46	2.10	1.45	1.53	85.98	92.21
NC_010481.6155782	*msrE*	1476	39.97	1.57	2.59	1.24	1.65	71.34	91.80
NC_011595.7072242	*tuf*	705	45.11	0.25	1.00	0.43	1.14	24.82	51.49
CP002695.1.gene10	*tuf*	1191	62.64	0.74	1.24	1.22	1.04	43.07	75.82
NC_002516.2.882884	*mexF*	3189	65.57	0.52	0.21	1.06	0.47	27.34	22.42
CP000647.1.gene3761	*tuf*	1185	54.09	0.64	0.54	1.38	0.72	18.65	40.68
NC_002516.2.881718	*tufA*	1194	59.63	0.48	0.53	1.01	0.97	22.11	28.06
CP003248.2.gene711	*rpsL*	375	62.67	-	-	-	-	ND	ND
CP003022.1.gene337	*msrE*	1476	39.97	1.72	1.37	1.79	1.07	64.16	75.07
X58272.1.gene1	bla_{OXA-5}	804	40.17	0.22	0.60	0.42	0.69	22.14	49.13

Table A.10. Average depth and gene similarity with BWA-MEM alignment for ARGs with ≥ 90% coverage in after disinfection (AD) samples for East Lansing (CAS) and Traverse City (MBR) WWTPs. ARGs from both CARD and ARG-ANNOT reference database are integrated into one data set.

Database Accession No.	Gene	Length	GC%	Avg Depth (x)		Depth Std		Gene Similarity (%)	
				ELAD	TCAD	ELAD	TCAD	ELAD	TCAD
(Sul)SulI:AF071413:6700-7539:840	*sul1*	855	60.58	2.66	2.25	1.32	0.87	97.19	98.33
AM087411.1.gene3	*aadA6*	846	53.66	3.51	2.56	4.25	1.19	100	100
AY743590.gene	*tet39*	1188	40.57	1.12	0.82	0.86	0.83	82.15	55.96
CP002695.1.gene3614	*tuf*	1191	62.64	24.15	47.95	14.13	26.90	100	100
EF102240.1.gene3	*msrE*	1476	39.97	1.64	1.01	1.18	2.31	78.46	41.02
(MLS)MphE:JF769133:8777-9661:885	*mphE*	900	35.78	1.55	2.61	0.70	1.07	92.56	98.25
(MLS)MsrE:JF769133:7246-8721:1476	*msrE*	1501	39.31	1.34	3.05	1.20	1.31	71.55	96.67
JQ364968.1.gene6	bla_{RTG-5}	378	43.12	1.36	1.89	0.50	1.17	99.21	93.90
NC_002516.2.881697	*tufB*	1194	59.63	24.97	34.68	14.62	24.84	96.40	92.55
NC_007682.3.4246769	*msrE*	1476	39.97	1.14	1.89	0.64	1.17	86.79	93.90
NC_010481.6155782	*msrE*	1476	39.97	1.40	0.82	0.75	0.83	91.26	55.96
NC_011595.7072242	*tuf*	705	45.11	10.41	4.54	10.43	4.52	100	90.78
CP002695.1.gene10	*tuf*	1191	62.64	21.20	45.58	13.25	26.59	98.74	100
NC_002516.2.882884	*mexF*	3189	65.57	10.59	13.75	11.47	13.03	94.45	91.66
CP000647.1.gene3761	*tuf*	1185	54.09	4.23	5.73	2.35	5.48	93.67	75.87
NC_002516.2.881718	*tufA*	1194	59.63	16.57	26.85	15.49	27.23	93.30	90.62
CP003248.2.gene711	*rpsL*	375	62.67	3.49	3.15	3.46	4.04	91.20	55.73
CP003022.1.gene337	*msrE*	1476	39.97	1.17	2.66	0.84	1.53	78.66	96.27
X58272.1.gene1	bla_{OXA-5}	804	40.17	0.22	0.60	0.42	0.65	22.14	91.17

REFERENCES

REFERENCES

Aires, J. R., Pechere, J. C., Van Delden, C., & Köhler, T. (2002). Amino acid residues essential for function of the MexF efflux pump protein of Pseudomonas aeruginosa. *Antimicrobial agents and chemotherapy*, *46*(7), 2169-2173.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. *Reference Source*.

Berglund, B., Fick, J., & Lindgren, P. E. (2015). Urban Wastewater Effluent Increases Antibiotic Resistance Gene Concentrations in a Receiving Northern European River. *Environmental Toxicology and Chemistry, 34*(1), 192-196.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, btu170.

Bondarczuk, K., Markowicz, A., & Piotrowska-Seget, Z. (2016). The urgent need for risk assessment on the antibiotic resistance spread via sewage sludge land application. *Environment International, 87,* 49-55.

Bonnin, R. A., Poirel, L., & Nordmann, P. (2012). A novel and hybrid composite transposon at the origin of acquisition of bla RTG-5 in Acinetobacter baumannii. *International journal of antimicrobial agents*, *40*(3), 257-259.

Bouki, C., Venieri, D., & Diamadopoulos, E. (2013). Detection and fate of antibiotic resistant bacteria in wastewater treatment plants: A review. *Ecotoxicology and Environmental Safety, 91*, 1-9.

Couture, F., Lachapelle, J., & Levesque, R. C. (1992). Phylogeny of LCR-1 and OXA-5 with class A and class D β-lactamases. *Molecular microbiology, 6*(12), 1693-1705.

Diehl, D. L., & LaPara, T. M. (2010). Effect of temperature on the fate of genes encoding tetracycline resistance and the integrase of class 1 integrons within anaerobic and aerobic digesters treating municipal wastewater solids. *Environmental science & technology, 44*(23), 9128-9133.

Everage, T. J., Boopathy, R., Nathaniel, R., LaFleur, G., & Doucet, J. (2014). A survey of antibiotic-resistant bacteria in a sewage treatment plant in Thibodaux, Louisiana, USA. *International Biodeterioration & Biodegradation, 95*, 2-10.

Fonseca, J. D., Knight, G. M., & McHugh, T. D. (2015). The complex evolution of antibiotic resistance in Mycobacterium tuberculosis. *International Journal of Infectious Diseases*, *32*, 94-100.

Finken, M., Kirschner, P., Meier, A., Wrede, A., & Böttger, E. C. (1993). Molecular basis of streptomycin resistance in Mycobacterium tuberculosis: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. *Molecular microbiology*, *9*(6), 1239-1246.

Gao, P., Mao, D., Luo, Y., Wang, L., Xu, B., & Xu, L. (2012a). Occurrence of sulfonamide and tetracycline-resistant bacteria and resistance genes in aquaculture environment. *Water Research*, *46*(7), 2355-2364.

Gao, P., Munir, M., & Xagoraraki, I. (2012b). Correlation of tetracycline and sulfonamide antibiotics with corresponding resistance genes and resistant bacteria in a conventional municipal wastewater treatment plant. *Science of the Total Environment*, *421*, 173-183.

Gomez-Alvarez, V., Revetta, R. P., & Santo Domingo, J. W. (2012). Metagenomic analyses of drinking water receiving different disinfection treatments. *Applied and environmental microbiology*, 78(17), 6095-6102.

Gonzalez-Villoria, A. M., & Valverde-Garduno, V. (2016). Antibiotic-Resistant Acinetobacter baumannii Increasing Success Remains a Challenge as a Nosocomial Pathogen. *Journal*

Guo, M. T., Yuan, Q. B., & Yang, J. (2015). Insights into the amplification of bacterial resistance to erythromycin in activated sludge. *Chemosphere*, *136*, 79-85.

Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L., & Rolain, J. M. (2014). ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrobial agents and chemotherapy*, *58*(1), 212-220.

Hall, C. C., Watkins, J. D., & Georgopapadakou, N. H. (1989). Effects of elfamycins on elongation factor Tu from Escherichia coli and Staphylococcus aureus. *Antimicrobial agents and chemotherapy*, *33*(3), 322-325.

Honore, N., & Cole, S. T. (1994). Streptomycin resistance in mycobacteria. *Antimicrobial agents and chemotherapy*, *38*(2), 238-242.

Huang, J. J., Hu, H. Y., Tang, F., Li, Y., Lu, S. Q., & Lu, Y. (2011). Inactivation and reactivation of antibiotic-resistant bacteria by chlorination in secondary effluents of a municipal wastewater treatment plant. *Water Research*, *45*(9), 2775-2781.

Kim, S., Park, H., & Chandran, K. (2010). Propensity of activated sludge to amplify or attenuate tetracycline resistance genes and tetracycline resistant bacteria: A mathematical modeling approach. *Chemosphere*, *78*(9), 1071-1077.

Kristiansson, E., Fick, J., Janzon, A., Grabic, R., Rutgersson, C., Weijdegård, B., ... & Larsson, D. J. (2011). Pyrosequencing of antibiotic-contaminated river sediments reveals high levels of resistance and gene transfer elements. *PloS one*, *6*(2), e17038.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, *9*(4), 357-359.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Li, H., & Homer, N. (2010). A survey of sequence alignment algorithms for next-generation sequencing. *Briefings in bioinformatics*, *11*(5), 473-483.

McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., ... & Kalan, L. (2013). The comprehensive antibiotic resistance database. *Antimicrobial agents and chemotherapy*, *57*(7), 3348-3357.

Manaia, C. M., Macedo, G., Fatta-Kassinos, D., & Nunes, O. C. (2015). Antibiotic resistance in urban aquatic environments: can it be controlled? *Applied microbiology and biotechnology*, 1-15.

Mao, D., Yu, S., Rysz, M., Luo, Y., Yang, F., Li, F., ... & Alvarez, P. J. J. (2015). Prevalence and proliferation of antibiotic resistance genes in two municipal wastewater treatment plants. *Water research*, *85*, 458-466.

Miele, A., Goldstein, B. P., Bandera, M., Jarvis, C., Resconi, A., & Williams, R. J. (1994). Differential Susceptibilities of Enterococcal Species to Elfamycin Antibiotics. *Journal of Clinical Microbiology*, *32*(8), 2016-2018.

Molnar, M., & Ilie, L. (2014). Correcting illumina data. *Briefings in bioinformatics*, bbu029.

Munir, M., Wong, K., & Xagoraraki, I. (2011). Release of antibiotic resistant bacteria and genes in the effluent and biosolids of five wastewater utilities in Michigan. *Water research*, *45*(2), 681-693.

Murray, G. E., Tobin, R. S., Junkins, B., & Kushner, D. J. (1984). Effect of chlorination on antibiotic resistance profiles of sewage-related bacteria. *Applied and environmental microbiology*, *48*(1), 73-77.

Novo, A., & Manaia, C. M. (2010). Factors influencing antibiotic resistance burden in municipal wastewater treatment plants. *Applied Microbiology and Biotechnology*, *87*(3), 1157-1166.

Olson, N. D., Lund, S. P., Colman, R. E., Foster, J. T., Sahl, J. W., Schupp, J. M., ... & Zook, J. M. (2015). Best practices for evaluating single nucleotide variant calling methods for microbial genomics. *Frontiers in genetics*, *6*.

Papadovasilaki, M., Oberthür, D., Gessmann, R., Sarrou, I., Betzel, C., Scoulica, E., & Petratos, K. (2015). Biophysical and enzymatic properties of aminoglycoside adenylyltransferase AadA6 from Pseudomonas aeruginosa. *Biochemistry and Biophysics Reports*, *4*, 152-157.

Penders, J., Stobberingh, E. E., Savelkoul, P. H., & Wolffs, P. F. (2013). The human microbiome as a reservoir of antimicrobial resistance. *Front Microbiol*, *4*(87), 542-549.

Rose, S., Desmolaize, B., Jaju, P., Wilhelm, C., Warrass, R., & Douthwaite, S. (2012). Multiplex PCR to identify macrolide resistance determinants in Mannheimia haemolytica and Pasteurella multocida. *Antimicrobial agents and chemotherapy*, *56*(7), 3664-3669.

Ruffalo, M., LaFramboise, T., & Koyutürk, M. (2011). Comparative analysis of algorithms for next-generation sequencing read alignment. *Bioinformatics*, *27*(20), 2790-2796.

Schmieder, R., & Edwards, R. (2012). Insights into antibiotic resistance through metagenomic approaches. *Future microbiology*, *7*(1), 73-89.

Sellyei, B., Varga, Z., Szentesi-Samu, K., Kaszanyitzky, E., & Magyar, T. (2009). Antimicrobial Susceptibility of Pasteurella Multocida Isolated from Swine and Poultry. *Acta Veterinaria Hungarica*, *57*(3), 357-367.

Shi, P., Jia, S., Zhang, X. X., Zhang, T., Cheng, S., & Li, A. (2013). Metagenomic insights into chlorination effects on microbial antibiotic resistance in drinking water. *Water research*, *47*(1), 111-120.

Sidrach-Cardona, R., Hijosa-Valsero, M., Marti, E., Balcazar, J. L., & Becares, E. (2014). Prevalence of antibiotic-resistant fecal bacteria in a river impacted by both an antibiotic production plant and urban treated discharges. *Science of the Total Environment*, *488*, 220-227.

Sims, D., Sudbery, I., Ilott, N. E., Heger, A., & Ponting, C. P. (2014). Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics*, *15*(2), 121-132.

Slekovec, C., Plantin, J., Cholley, P., Thouverez, M., Talon, D., Bertrand, X., et al. (2012). Tracking Down Antibiotic-Resistant Pseudomonas aeruginosa Isolates in a Wastewater Network. *Plos One*, *7*(12).

Sottani, C., Islam, K., Soffientini, A., Zerilli, L. F., & Seraglia, R. (1993). Studies on the Interaction of Elfamycin Antibiotics with Elongation Factor-Tu by Mass Spectroscopic Techniques. *Rapid Communications in Mass Spectrometry*, *7*(7), 680-683.

Szczepanowski, R., Linke, B., Krahn, I., Gartemann, K. H., Guetzkow, T., Eichler, W., ... & Schlueter, A. (2009). Detection of 140 clinically relevant antibiotic-resistance genes in the plasmid metagenome of wastewater treatment plant bacteria showing reduced susceptibility to selected antibiotics. *Microbiology*, *155*(7), 2306-2319.

Varela, A. R., & Manaia, C. M. (2013). Human health implications of clinically relevant bacteria in wastewater habitats. *Environmental Science and Pollution Research*, *20*(6), 3550-3569.

Vaz-Moreira, I., Varela, A. R., Pereira, T. V., Fochat, R. C., & Manaia, C. M. (2015). Multidrug Resistance in quinolone-resistant gram-negative bacteria isolated from hospital effluent and the municipal wastewater treatment plant. *Microbial Drug Resistance*.

Weber, D. J., Wolfson, J. S., Swartz, M. N., & Hooper, D. C. (1984). Pasteurella-Multocida Infections - Report of 34 Cases and Review of the Literature. *Medicine, 63*(3), 133-154.

Wright, G. D. (1999). Aminoglycoside-modifying enzymes. *Current opinion in microbiology, 2*(5), 499-503.

Xia, J., Gao, J., & Tang, W. (2016). Nosocomial infection and its molecular mechanisms of antibiotic resistance. *Bioscience trends*, (0).

Yuan, Q. B., Guo, M. T., & Yang, J. (2015). Fate of Antibiotic Resistant Bacteria and Genes during Wastewater Chlorination: Implication for Antibiotic Resistance Control. *Plos One, 10*(3).

Zarrilli, R., Vitale, D., Di Popolo, A., Bagattini, M., Daoud, Z., Khan, A. U., ... & Triassi, M. (2008). A plasmid-borne blaOXA-58 gene confers imipenem resistance to Acinetobacter baumannii isolates from a Lebanese hospital. *Antimicrobial agents and chemotherapy, 52*(11), 4115-4120.

Zhang, C. M., Du, C., Xu, H., Miao, Y. H., Cheng, Y. Y., Tang, H., et al. (2015a). Occurrence of tetracycline-resistant fecal coliforms and their resistance genes in an urban river impacted by municipal wastewater treatment plant discharges. *Journal of Environmental Science and Health Part a-Toxic/Hazardous Substances & Environmental Engineering, 50*(7), 744-749.

Zhang, S. H., Han, B., Gu, J., Wang, C., Wang, P. F., Ma, Y. Y., et al. (2015b). Fate of antibiotic resistant cultivable heterotrophic bacteria and antibiotic resistance genes in wastewater treatment processes. *Chemosphere, 135*, 138-145.

Zhang, T., Yang, Y., & Pruden, A. (2015c). Effect of temperature on removal of antibiotic resistance genes by anaerobic digestion of activated sludge revealed by metagenomic approach. *Applied microbiology and biotechnology, 99*(18), 7771-7779.

Zhang, X., Wu, B., Zhang, Y., Zhang, T., Yang, L., Fang, H. H., ... & Cheng, S. (2009a). Class 1 integronase gene and tetracycline resistance genes tetA and tetC in different water environments of Jiangsu Province, China. *Ecotoxicology, 18*(6), 652-660.

Zhang, X. X., & Zhang, T. (2011). Occurrence, abundance, and diversity of tetracycline resistance genes in 15 sewage treatment plants across China and other global locations. *Environmental science & technology, 45*(7), 2598-2604.

Zhang, X. X., Zhang, T., & Fang, H. H. (2009b). Antibiotic resistance genes in water environment. *Applied microbiology and biotechnology, 82*(3), 397-414.